

KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
 KW immunomodulator; cytostatic; neurotropic; neuroprotective;
 KW antiparkinsonian; antilipase; gene therapy; human disease;
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; dyslipidaemia.
 XX
 OS Homo sapiens.
 PN WO2003029424-A2.
 PD 10-APR-2003.
 XX
 PF 02-OCT-2002; 2002WO-US031373.
 XX
 PR 02-OCT-2001; 2001US-0326483P.
 PR 05-OCT-2001; 2001US-0327435P.
 PR 05-OCT-2001; 2001US-0327449P.
 PR 09-OCT-2001; 2001US-0327917P.
 PR 09-OCT-2001; 2001US-0328029P.
 PR 09-OCT-2001; 2001US-0328044P.
 PR 09-OCT-2001; 2001US-0328056P.
 PR 12-OCT-2001; 2001US-0328849P.
 PR 15-OCT-2001; 2001US-0329414P.
 PR 17-OCT-2001; 2001US-0330142P.
 PR 18-OCT-2001; 2001US-0330309P.
 PR 22-OCT-2001; 2001US-0341058P.
 PR 24-OCT-2001; 2001US-0343266P.
 PR 24-OCT-2001; 2001US-0343629P.
 PR 29-OCT-2001; 2001US-0349575P.
 PR 01-NOV-2001; 2001US-0346357P.
 PR 17-APR-2002; 2002US-0373260P.
 PR 19-APR-2002; 2002US-0373815P.
 PR 19-APR-2002; 2002US-0373817P.
 PR 19-APR-2002; 2002US-0373826P.
 PR 19-APR-2002; 2002US-0373884P.
 PR 22-APR-2002; 2002US-0374977P.
 PR 16-MAY-2002; 2002US-0381037P.
 PR 16-MAY-2002; 2002US-0381038P.
 PR 16-MAY-2002; 2002US-0381042P.
 PR 17-MAY-2002; 2002US-0381642P.
 PR 28-MAY-2002; 2002US-0383656P.
 PR 29-MAY-2002; 2002US-0383831P.
 PR 25-JUN-2002; 2002US-0393335P.
 PR 01-OCT-2002; 2002US-00262511.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
 PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malvankar UM;
 PI Ott T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shency SG;
 PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dippio VA;
 PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;
 XX
 WPI; 2003-381626/36.
 DR N-PSDB; ADA05733.
 XX
 XX New NOVX polypeptides and nucleic acids, useful for diagnosing,
 PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
 PT cancer or dyslipidaemia, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.
 XX
 PS Claim 1; Page 170; 586pp; English.
 XX
 CC The present invention describes NOVX proteins, where X can be 1 to 55
 CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide
 CC described above and a carrier; (2) a kit comprising, in one or more
 CC containers, the composition described above; (3) an isolated nucleic acid
 CC molecule which encodes a NOVX protein of the invention; (4) a vector
 CC comprising the nucleic acid molecule described above; (5) a cell
 CC comprising the above vector; (6) an antibody that immunospecifically
 CC binds to the polypeptide described above; (7) methods for determining the

CC presence or amount of the above polypeptide or nucleic acid molecule in a
 CC sample; (8) methods for determining the presence of or predisposition to
 CC a disease associated with altered levels of expression of the above
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
 CC method of identifying an agent that binds to the polypeptide described
 CC above; (10) a method for identifying a potential therapeutic agent for
 CC use in treating a pathology that is related to an aberrant expression or
 CC aberrant physiological interactions of the polypeptide; (11) a method of
 CC screening for a modulator of activity or of latency or predisposition to
 CC a pathology associated with the polypeptide; (12) a method for modulating
 CC the activity of the polypeptide associated with the above polypeptide in a
 CC or preventing a pathology associated with the above polypeptide in a
 CC mammal; and (14) a method for producing the above polypeptide. NOVX
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,
 CC immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian
 CC and antilipase activities, and can be used in gene therapy. The
 CC polypeptide is useful in manufacturing a medicament for treating a
 CC syndrome associated with a human disease. The polypeptide or the nucleic
 CC acid molecule may be used to diagnose, treat or prevent metabolic
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
 CC disease, immune disorders, haematopoietic disorders and various
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation
 CC probes, in chromosome mapping, tissue typing, preventive medicine and
 CC pharmacogenomics. The present sequence represents a human NOVX from the
 CC present invention.
 XX
 SQ Sequence 252 AA;
 Query Match 100.0%; Score 40; DB 6; Length 252;
 Best Local Similarity 100.0%; Pred. No. 2.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 RLSSMVKKV 9
 Db 122 RLSSMVKKV 130
 RESULT 7
 AAR67888
 ID AAR67888 standard; protein; 253 AA.
 AC AAR67888;
 XX
 XX 25-MAR-2003 (revised)
 DT 09-AUG-1995 (first entry)
 DE Human stratum corneum chymotrophic recombinant enzyme (SCCE).
 XX
 XX Stratum corneum chymotryptic enzyme; skin disorder; acne; psoriasis;
 KW callosities; keratosis pilaris; ichthyoses; eczema.
 XX
 OS Homo sapiens.
 XX
 PN WO9500651-A1.
 XX
 XX 05-JAN-1995.
 PD
 PF 20-JUN-1994; 94WO-IB000166.
 XX
 PR 18-JUN-1993; 93DK-00000725.
 XX
 PA (SYMB-) SYMBICOM AB.
 XX
 PI Egelrud T, Hansson L;
 XX
 WPI; 1995-052088/07.
 DR N-PSDB; AAQ81203.
 XX
 PT Nucleotide sequences encoding stratum corneum chymotryptic enzyme - and
 PT related vectors, transformed cells and polypeptides, useful for treating
 PT skin disorders, e.g. acne or psoriasis, and for identification of
 PT specific inhibitors.

XX Disclosure; Page 97; 137pp; English.

XX The enzyme encoded by this sequence is used in pharmaceutical, cosmetic

CC and skin care products, especially to treat and prevent acne, xeroderma,

CC or other hyperkeratotic conditions (e.g. callusities or keratosis

CC plaris), ichthyoses, psoriasis, eczema, etc. It is produced

CC recombinantly following mammal, insect, plant, or microorganism

CC transformation with plasmid pS507. (Updated on 25-MAR-2003 to correct PN

CC field.)

XX Sequence 253 AA;

SQ

Query Match 100.0%; Score 40; DB 2; Length 253;

Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSWVKV 9

DB 123 RLSSWVKV 131

RESULT 8

AAW05383

ID AAW05383 standard; protein; 253 AA.

AC AAW05383;

XX 31-DEC-1996 (first entry)

XX Human amyloid precursor protein protease.

XX Amyloid precursor protein protease; Alzheimer's disease; diagnosis;

XX therapy.

XX Homo sapiens.

XX WO9631122-A1.

XX 10-OCT-1996.

XX 02-APR-1996; 96WO-US004294.

XX 04-APR-1995; 95US-00416257.

XX (ELIL) LILLY & CO ELI.

XX Dixon EP, Johnstone EM, Little SP;

XX WPI; 1996-464694/46.

XX N-PSDB; AAT39783.

XX New isolated human amyloid precursor protein protease - used to develop

PT prods. for the treatment or diagnosis of associated conditions, esp.

PT Alzheimer's disease.

XX Claim 1; Page 44-45; 55pp; English.

XX Human amyloid precursor protein protease (AAW05383) is involved in the

CC processing or clearance of amyloid precursor protein to form beta-amyloid

CC peptide. Its amino acid sequence was deduced from a cDNA clone (AAT39783)

CC obtd. from a human lung library. Recombinant protease can be produced in

CC transformed or transfected prokaryotic (partic. E. coli) or eukaryotic

CC (partic. AV-120 host cells. It is used to develop products for the design

CC and testing of cpds. useful for treating or preventing conditions

CC associated with beta-amyloid peptide, esp. Alzheimer's disease

XX Sequence 253 AA;

SQ

Query Match 100.0%; Score 40; DB 2; Length 253;

Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSWVKV 9

DB 123 RLSSWVKV 131

RESULT 9

ABB84421

ID ABB84421 standard; peptide; 253 AA.

XX ABB84421;

XX 08-NOV-2002 (first entry)

XX Human SCCE protein N-terminal fragment SEQ ID 48.

DE SCCE; human; stratum corneum chymotryptic enzyme; kallikrein 7;

XX serine protease; transgenic mammal; skin; skin disease; skin cancer;

XX hyperkeratosis; acanthosis; epidermal inflammation; dermal inflammation;

XX pruritus; atopic dermatitis; eczema; acne; itch; KLK7; enzyme.

XX Homo sapiens.

XX WO200262135-A2.

XX 15-AUG-2002.

XX 08-FEB-2002; 2002WO-IB001300.

XX 09-FEB-2001; 2001CA-02332655.

XX 09-FEB-2001; 2001DK-00000218.

XX (EGEL/) EGELRUD T.

XX (HANS/) HANSSON L.

XX Egelrud T, Hansson L;

XX WPI; 2002-643380/69.

XX Transgenic mammal or its embryo useful as model for human disease, has

PT heterologous nucleotide sequence coding for stratum corneum chymotryptic

PT enzyme operably linked to promoter that drives its expression in skin.

XX Example 6; Page 37; 74pp; English.

XX This invention describes a novel non-human transgenic mammal or mammalian

CC embryo having integrated within its genome, a heterologous nucleotide

CC sequence comprising at least a significant part of a nucleotide sequence

CC coding for a stratum corneum chymotryptic enzyme (SCCE) or its variant,

CC operably linked to a promoter that drives expression of heterologous scce

CC or its variant in skin. The product of the invention is useful as a model

CC for the study of disease with the aim of improving treatment, to relieve

CC or ameliorate a pathogenic condition, for development or testing of a

CC cosmetic or a pharmaceutical formulation, and for the development of a

CC diagnostic method. It can also be used as a model for a skin disease or

CC skin cancer. The invention is also useful for screening or identifying a

CC compound or composition effective for the prevention or treatment of an

CC abnormal or unwanted phenotype, and for screening or identifying a

CC compound or composition effective for the prevention or treatment of

CC inflammatory skin diseases selected from diseases consisting of epidermal

CC hyperkeratosis, acanthosis, epidermal inflammation, dermal inflammation,

CC pruritus, atopic dermatitis, eczema, acne and inherited skin diseases

CC with epidermal hyperkeratosis. The mammal of the invention is also useful

CC as a model for further studies of itch mechanisms and the testing of

CC potential compounds and compositions for relieve of various skin diseases

CC where itch is a component. This sequence represents the N-terminal

CC fragment of the human stratum corneum chymotryptic enzyme, SCCE

CC synonymous with human kallikrein 7 (KLK7), used in the development of the

CC transgenic mammals described in the invention

XX Sequence 253 AA;

Query Match 100.0%; Score 40; DB 5; Length 253;

Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMWKVK 9
 123 RLSSMWKVK 131

Db

RESULT 10
 ABB84406
 ID ABB84406 standard; protein; 253 AA.

XX AC ABB84406;
 XX DT 08-NOV-2002 (first entry)
 XX DE Human SCCE protein.

XX KW SCCE; human; stratum corneum chymotryptic enzyme; kallikrein 7;
 KW serine protease; transgenic mammal; skin; skin disease; skin cancer;
 KW hyperkeratosis; acanthosis; epidermal inflammation; dermal inflammation;
 KW pruritus; atopic dermatitis; eczema; acne; itch; K&K7; enzyme.

XX OS Homo sapiens.
 XX PN WO200262135-A2.
 XX PD 15-AUG-2002.
 XX PF 08-FEB-2002; 2002WO-IB001300.
 XX PR 09-FEB-2001; 2001CA-02332655.
 XX PR 09-FEB-2001; 2001DK-00000218.
 XX PA (EGL/) EGELRUD T.
 XX PA (HANS/) HANSSON L.
 XX PI Egelrud T, Hansson L;
 XX DR WPI; 2002-643380/69.
 XX DR N-PSDB; ABQ76226.

XX Transgenic mammal or its embryo useful as model for human disease, has
 PT heterologous nucleotide sequence coding for stratum corneum chymotryptic
 PT enzyme operably linked to promoter that drives its expression in skin.

XX Claim 10; Page 58-59; 74pp; English.

XX This invention describes a novel non-human transgenic mammal or mammalian
 CC embryo having integrated within its genome, a heterologous nucleotide
 CC sequence comprising at least a significant part of a nucleotide sequence
 CC coding for a stratum corneum chymotryptic enzyme (SCCE) or its variant,
 CC operably linked to a promoter that drives expression of heterologous scce
 CC or its variant in skin. The product of the invention is useful as a model
 CC for the study of disease with the aim of improving treatment, to relieve
 CC or ameliorate a pathogenic condition, for development or testing of a
 CC cosmetic or a pharmaceutical formulation, and for the development of a
 CC diagnostic method. It can also be used as a model for a skin disease or
 CC skin cancer. The invention is also useful for screening or identifying a
 CC compound or composition effective for the prevention or treatment of an
 CC abnormal or unwanted phenotype, and for screening or identifying a
 CC compound or composition effective for the prevention or treatment of
 CC inflammatory skin diseases selected from diseases consisting of epidermal
 CC inflammation, epidermal inflammation, dermal inflammation,
 CC hyperkeratosis, acanthosis, eczema, acne and inherited skin diseases
 CC pruritus, atopic dermatitis, epidermal inflammation, dermal inflammation,
 CC as a model for further studies of itch mechanisms and the testing of
 CC potential compounds and compositions for relieve of various skin diseases
 CC where itch is a component. This sequence represents the human stratum
 CC corneum chymotryptic enzyme, SCCE which is a serine protease synonymous
 CC with human kallikrein 7 (KLK7) and is used in the development of the
 CC transgenic mammals described in the invention

XX Sequence 253 AA;

Query Match 100.0%; Score 40; DB 5; Length 253;
 Best Local Similarity 100.0%; Pred. No. 2.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMWKVK 9
 123 RLSSMWKVK 131

Db

RESULT 11
 AAU82740
 ID AAU82740 standard; protein; 253 AA.

XX AC AAU82740;
 XX DT 23-APR-2002 (first entry)

XX Amino acid sequence of novel human protease #39.

XX Human; protease; cancer; immune-related disorder; cardiovascular disease;
 KW neuronal-associated disease; metabolic disorder; inflammatory disorder;
 KW nervous system disorder; sexual dysfunction; pain; mood disorder;
 KW hypertension; psychotic disorder; neurological disorder; dyskinesia;
 KW viral infection; human immunodeficiency virus; HIV; non-viral infection;
 KW ocular disease; cytostatic; enzyme.

XX OS Homo sapiens.
 XX PN WO200200860-A2.
 XX PD 03-JAN-2002.
 XX PF 26-JUN-2001; 2001WO-US020171.
 XX PR 26-JUN-2000; 2000US-0214047P.
 XX PA (SUGE-) SUGEN INC.
 XX PI Plowman G, Whyte D, Sudarsanam S, Manning G, Caenepeel S;
 XX PI Charyczak G;
 XX WPI; 2002-139913/18.
 XX DR N-PSDB; ABK31782.

XX Nucleic acids encoding novel human proteases, useful for useful for
 PT treating diseases and disorders such as cancers, immune-related diseases
 PT and disorders, cardiovascular disease (e.g. restenosis) and inflammatory
 PT disorders.

XX Claim 6; Fig 2N; 313pp; English.

XX The present invention relates to the isolation of novel human proteases,
 CC and the nucleic acids encoding them. The sequences of the invention are
 CC useful for treating diseases and disorders such as cancers (e.g. breast,
 CC colon, lung), immune-related diseases and disorders (e.g. inflammatory
 CC diseases and asthma), cardiovascular diseases (e.g. restenosis and
 CC coronary thrombosis), brain or neuronal-associated diseases, metabolic
 CC disorders (e.g. diabetes, obesity), inflammatory disorders (e.g.
 CC rheumatoid arthritis and psoriasis), central or peripheral nervous system
 CC diseases, migraines, pain, sexual dysfunction, mood disorders, attention
 CC disorders, cognitive disorders, hypertension, mood disorders, psychotic
 CC disorders, neurological disorders (e.g. Alzheimer's disease, Parkinson's
 CC disease), and dyskinesias. The nucleic acids and polypeptides are also
 CC useful for treating viral infections caused by human immunodeficiency
 CC virus (HIV), and non-viral infections such as ocular disease (e.g.
 CC glaucoma) and macular degeneration. AAU82702-AAU82760 represent the novel
 CC human proteases of the invention

XX Sequence 253 AA;

Query Match 100.0%; Score 40; DB 5; Length 253;
 Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKV 9
 |||||
 Db 123 RLSSMVKV 131

RESULT 12
 ABU07440
 ID ABU07440 standard; protein; 253 AA.
 AC ABU07440;
 XX
 XX 28-JAN-2003 (first entry)
 DT
 DE Protein differentially regulated in prostate cancer #43.
 KW Prostate cancer; gene expression; differential regulation;
 KW molecular marker; drug target; cancer detection; cancer diagnosis;
 KW cancer staging; cancer grading; cancer assessing; cancer monitoring.
 XX
 OS Homo sapiens.
 XX
 XX WO200281638-A2.
 PN
 XX
 PD 17-OCT-2002.
 XX
 XX 08-APR-2002; 2002WO-US010824.
 PF
 XX 06-APR-2001; 2001US-0281731P.
 PR 06-APR-2001; 2001US-0281732P.
 XX
 XX (ORIG-) ORIGENE TECHNOLOGIES INC.
 PA
 XX
 XX Sun Z, Jay G;
 FI
 XX WPI; 2003-058520/05.
 DR N-PSDB; ABX10343.
 XX
 XX Novel genes which are differentially regulated in prostate cancer, useful
 PT for diagnosing prostate cancer in prostate tissue sample and assessing
 PT therapeutic or preventive intervention in prostate cancer patients.
 XX
 PS Claim 1; Page 293-294; 416pp; English.

The invention describes genes (I) which are differentially regulated in prostate cancer. (I) is useful for diagnosing a prostate cancer in a sample comprising prostate tissue, which involves determining the number of target genes which are differentially-regulated in the sample, where the number is indicative of the probability that the sample comprises prostate cancer. (I) is useful for assessing a therapeutic or preventive intervention in a subject having a prostate cancer, which involves determining the expression levels in a sample comprising prostate tissue of target genes which are differentially-regulated in prostate cancer. Preferably, the expression levels of at least 10 genes are determined. (I) is also useful for identifying agents that modulate a biological activity of a polypeptide differentially-regulated in prostate cancer cells, which involves contacting a polypeptide differentially-regulated in prostate cancer cells with a test agent under conditions effective for determining whether the test agent modulates the biological activity. (I) is useful as molecular markers, as drug targets, and for detecting, diagnosing, staging, grading, assessing, monitoring, prognosticating, preventing or treating, determining predisposition to diseases and conditions especially relating to prostate cancer. (I) and its expression products are used in the diagnostic test to assay for presence of cancer e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in blood etc. (I) is useful for assessing cancer e.g., to determine the type of cancer, its stage of development, the nature of genetic defect, etc. The polypeptide encoded by (I) can be used as target for therapy or drug discovery. (I) can also be used for expressing the polypeptide and thus for searching specific binding partners of the polypeptide. (I) is useful in therapeutic applications to treat prostate cancer. The identification

CC of specific genes, and groups of genes, expressed in pathways
 CC physiologically relevant to prostate cancer permits the definition of
 CC functional and disease pathways and the delineation of targets in these
 CC pathways which are useful in diagnostic, therapeutic, and clinical
 CC applications. This is the amino acid sequence of a protein differentially
 CC regulated in prostate cancer
 XX

XX Sequence 253 AA;

Query Match 100.0%; Score 40; DB 6; Length 253;
 Best Local Similarity 100.0%; Pred. No. 2.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKV 9
 |||||
 Db 123 RLSSMVKV 131

RESULT 13
 ABU07471
 ID ABU07471 standard; protein; 253 AA.
 AC ABU07471;
 XX
 XX 28-JAN-2003 (first entry)
 DT
 DE Protein differentially regulated in prostate cancer #74.
 XX
 KW Prostate cancer; gene expression; differential regulation;
 KW molecular marker; drug target; cancer detection; cancer diagnosis;
 KW cancer staging; cancer grading; cancer assessing; cancer monitoring.
 XX
 OS Homo sapiens.
 XX
 XX WO200281638-A2.
 PN
 XX
 PD 17-OCT-2002.
 XX
 XX 08-APR-2002; 2002WO-US010824.
 PF
 XX 06-APR-2001; 2001US-0281731P.
 PR 06-APR-2001; 2001US-0281732P.
 XX
 XX (ORIG-) ORIGENE TECHNOLOGIES INC.
 PA
 XX
 XX Sun Z, Jay G;
 FI
 XX WPI; 2003-058520/05.
 DR N-PSDB; ABX10375.
 XX

Novel genes which are differentially regulated in prostate cancer, useful
 PT for diagnosing prostate cancer in prostate tissue sample and assessing
 PT therapeutic or preventive intervention in prostate cancer patients.
 XX

PS Claim 1; Page 351; 416pp; English.

The invention describes genes (I) which are differentially regulated in prostate cancer. (I) is useful for diagnosing a prostate cancer in a sample comprising prostate tissue, which involves determining the number of target genes which are differentially-regulated in the sample, where the number is indicative of the probability that the sample comprises prostate cancer. (I) is useful for assessing a therapeutic or preventive intervention in a subject having a prostate cancer, which involves determining the expression levels in a sample comprising prostate tissue of target genes which are differentially-regulated in prostate cancer. Preferably, the expression levels of at least 10 genes are determined. (I) is also useful for identifying agents that modulate a biological activity of a polypeptide differentially-regulated in prostate cancer cells, which involves contacting a polypeptide differentially-regulated in prostate cancer cells with a test agent under conditions effective for determining whether the test agent modulates the biological activity. (I) is useful as molecular markers, as drug targets, and for detecting,

CC diagnosing, staging, grading, assessing, monitoring, prognosticating,
 CC preventing or treating, determining predisposition to diseases and
 CC conditions especially relating to prostate cancer. (I) and its expression
 CC products are used in the diagnostic test to assay for presence of cancer
 CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
 CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
 CC of cancer, its stage of development, the nature of genetic defect, etc.
 CC The polypeptide encoded by (I) can be used as target for therapy or drug
 CC discovery. (I) can also be used for expressing the polypeptide and thus
 CC for searching specific binding partners of the polypeptide. (I) is useful
 CC in therapeutic applications to treat prostate cancer. The identification
 CC of specific genes, and groups of genes, expressed in pathways
 CC physiologically relevant to prostate cancer permits the definition of
 CC functional and disease pathways and the delineation of targets in these
 CC pathways which are useful in diagnostic, therapeutic, and clinical
 CC applications. This is the amino acid sequence of a protein differentially
 CC regulated in prostate cancer

XX SQ Sequence 253 AA;

Query Match 100.0%; Score 40; DB 6; Length 253;
 Best Local Similarity 100.0%; Pred. No. 2.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9
 |||||
 DB 123 RLSSMVKKV 131

RESULT 14

ABR58471 ABR58471 standard; protein; 253 AA.

XX AC ABR58471;

XX DT 07-JUL-2003 (first entry)

XX DE Human stratum corneum chymotryptic enzyme - ovarian cancer clone O1676P.

XX KW Human; cytostatic; Gene therapy; vaccine; cancer; ovarian cancer.

XX OS Homo sapiens.

XX EN WO2003029468-A1.

XX PD 10-APR-2003.

XX PF 02-OCT-2002; 2002WO-US031467.

XX PR 02-OCT-2001; 2001US-0327135P.

XX PR 30-MAY-2002; 2002US-0384531P.

XX PA (CORI-) CORIXA CORP.

XX PI Algate PA, Mannion J;

XX DR WPI; 2003-372001/35.

XX PT New polynucleotide and polypeptide useful for diagnosing and/or treating
 PT cancer, particularly ovarian cancer, and as a vaccine.

XX PS Claim 2; Page 157-158; 169pp; English.

XX CC The invention relates to a novel isolated polynucleotide. The
 CC polynucleotides of the invention have cytostatic activity, and may have a
 CC use in gene therapy, and in a vaccine. The composition and methods are
 CC useful in diagnosing and/or treating cancer, particularly ovarian cancer.
 CC The composition may also be used as a vaccine to prevent cancer. The
 CC present sequence is used in the exemplification of the invention

XX SQ Sequence 253 AA;

Query Match 100.0%; Score 40; DB 6; Length 253;

Best Local Similarity 100.0%; Pred. No. 2.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9
 |||||
 DB 123 RLSSMVKKV 131

RESULT 15

ADB80484 ADB80484 standard; protein; 253 AA.

XX AC ADB80484;

XX DT 04-DEC-2003 (first entry)

XX DE Ovarian cancer-associated protein #24.

XX KW cytostatic; Gene therapy; vaccine; ovarian cancer; diagnosis;
 KW post-operative chemotherapy; radiation therapy; tumour prognosis;
 KW pre-cancerous lesion detection.

XX OS Homo sapiens.

XX PN WO2002102235-A2.

XX PD 27-DEC-2002.

XX PF 18-JUN-2002; 2002WO-US019297.

XX PR 18-JUN-2001; 2001US-0299234P.

XX PR 27-AUG-2001; 2001US-0315287P.

XX PR 05-SEP-2001; 2001US-0317544P.

XX PR 13-NOV-2001; 2001US-0350666P.

XX PR 12-APR-2002; 2002US-0372246P.

XX PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX PI Mack DH, Gish KC;

XX DR WPI; 2003-167431/16.

XX DR N-PSDB; ADB80483.

XX PT Detecting an ovarian cancer-associated transcript in a cell from a
 PT patient, comprises contacting a biological sample from the patient with a
 PT polynucleotide that hybridizes to an ovarian cancer gene.

XX PS Claim 13; Page 291; 332pp; English.

XX CC The invention relates to a method of detecting an ovarian cancer-
 CC associated transcript in a cell from a patient, by contacting a
 CC biological sample from the patient with a polynucleotide that selectively
 CC hybridizes to a sequence, at least 80% identical to any of one of 80
 CC nucleic acid sequences given in the specification. The method is useful
 CC in diagnosing ovarian cancer and in identifying and using agents and/or
 CC targets that inhibit ovarian cancer. The nucleic acid molecule,
 CC polypeptide and the antibody may also be used in detecting ovarian
 CC cancers, monitoring and early detection of relapse following treatment,
 CC monitoring response to therapy, selecting patients for post-operative
 CC chemotherapy or radiation therapy, in selecting mode of therapy,
 CC determining tumour prognosis, early detection of pre-cancerous lesions,
 CC and as vaccines. This sequence corresponds to one of the proteins used
 CC for the detection method of the invention.

XX SQ Sequence 253 AA;

Query Match 100.0%; Score 40; DB 7; Length 253;

Best Local Similarity 100.0%; Pred. No. 2.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9
 |||||
 DB 123 RLSSMVKKV 131

Search completed: March 1, 2004, 17:28:47
Job time : 46.5556 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:35:01 ; Search time 24.1111 Seconds
(without alignments)
78.818 Million cell updates/sec

Title: US-09-905-083-32
Perfect score: 40
Sequence: 1 RLSSWVKV 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 21153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	9	9	US-09-918-243-32
2	40	100.0	9	9	Sequence 32, Appl
3	40	100.0	144	9	US-09-905-083-32
4	40	100.0	144	9	US-09-796-294-4
5	40	100.0	253	14	US-10-461-787-4
6	40	100.0	253	9	US-09-888-615-98
7	40	100.0	253	9	US-09-764-762-3
8	40	100.0	253	14	US-10-264-283-90
9	40	100.0	253	15	US-10-295-027-498
10	36	90.0	9	15	US-10-173-999-48
11	36	90.0	9	9	US-09-918-243-110
12	33	82.5	394	14	US-10-094-240-4
13	33	82.5	394	14	US-10-056-405-4
14	31	77.5	9	9	US-09-918-243-79
15	31	77.5	9	9	US-09-918-243-87

16	31	77.5	9	9	US-09-905-083-79	Sequence 79, Appl
17	31	77.5	9	9	US-09-905-083-87	Sequence 87, Appl
18	31	77.5	2159	14	US-10-208-948-24	Sequence 24, Appl
19	30	75.0	833	15	US-10-369-493-1587	Sequence 1587, Ap
20	30	75.0	1164	15	US-10-369-493-9275	Sequence 9275, Ap
21	30	75.0	1164	15	US-10-369-493-9432	Sequence 9432, Ap
22	30	75.0	1167	15	US-10-369-493-17743	Sequence 17743, A
23	30	75.0	1846	15	US-10-369-493-4974	Sequence 4974, Ap
24	29	72.5	436	14	US-10-156-761-10832	Sequence 10832, A
25	29	72.5	469	15	US-10-369-493-18267	Sequence 18267, A
26	28	70.0	141	14	US-10-028-386-29875	Sequence 29875, A
27	28	70.0	175	15	US-10-028-248A-112	Sequence 112, App
28	28	70.0	317	15	US-10-107-782-112	Sequence 112, App
29	28	70.0	317	15	US-10-369-493-23122	Sequence 23122, A
30	28	70.0	448	9	US-09-871-212-8	Sequence 8, Appli
31	28	70.0	459	9	US-09-815-242-12703	Sequence 12703, A
32	28	70.0	474	15	US-10-369-493-16413	Sequence 16413, A
33	28	70.0	485	14	US-10-309-629-2	Sequence 2, Appli
34	28	70.0	523	14	US-10-014-101-6	Sequence 6, Appli
35	28	70.0	625	15	US-10-369-493-17172	Sequence 17172, A
36	28	70.0	758	15	US-10-094-749-2451	Sequence 2451, Ap
37	28	70.0	804	10	US-09-116-676-10	Sequence 10, Appl
38	28	70.0	896	8	US-08-779-457-3	Sequence 3, Appli
39	28	70.0	896	13	US-10-095-929-10	Sequence 10, Appl
40	28	70.0	896	14	US-10-214-802-3	Sequence 3, Appli
41	28	70.0	898	14	US-10-245-616-3	Sequence 3, Appli
42	28	70.0	906	13	US-10-095-929-9	Sequence 9, Appli
43	28	70.0	923	8	US-08-779-457-4	Sequence 4, Appli
44	28	70.0	923	14	US-10-214-802-4	Sequence 4, Appli
45	28	70.0	958	13	US-10-095-929-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-918-243-32
; Sequence 32, Application US/09918243
; Patent No. US20020142317A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 32
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 123-131 of the SCCE protein
US-09-918-243-32

Query Match 100.0%; Score 40; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSWVKV 9
Db 1 RLSSWVKV 9

RESULT 2
US-09-905-083-32
; Sequence 32, Application US/09905083
; Patent No. US20020146708A1
; GENERAL INFORMATION:

```
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP/C/Div
; CURRENT APPLICATION NUMBER: US/09/905,083
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 09/502,600
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 32
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 123-131 of the SCCE protein
US-09-905-083-32

Query Match          100.0%; Score 40; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 RLSSMVKKV 9
Db      1 RLSSMVKKV 9

RESULT 3
US-09-796-294-4
; Sequence 4, Application US/09796234
; Patent No. US20020037581A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Underwood, Lowell J.
; TITLE OF INVENTION: Extracellular Serine Protease
; FILE REFERENCE: D6020CIP3
; CURRENT APPLICATION NUMBER: US/09/796,234
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/618,259
; PRIOR FILING DATE: 2000-07-18
; NUMBER OF SEQ ID NOS: 72
; SEQ ID NO 4
; LENGTH: 144
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; NAME/KEY: DOMAIN
; OTHER INFORMATION: Amino acid sequence of stratum corneum chymotryptic
; OTHER INFORMATION: enzyme (scce) catalytic domain
US-09-796-294-4

Query Match          100.0%; Score 40; DB 9; Length 144;
Best Local Similarity 100.0%; Pred. No. 0.97;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 RLSSMVKKV 9
Db      60 RLSSMVKKV 68

RESULT 4
US-10-461-787-4
; Sequence 4, Application US/10461787
; Publication No. US20030199010A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Underwood, Lowell J.
; TITLE OF INVENTION: No. US20030199010A1 Extracellular Serine Protease
; FILE REFERENCE: D6020CIP2
; CURRENT APPLICATION NUMBER: US/10/461,787
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 09/618,259
; PRIOR FILING DATE: 2000-07-18
```

```
; PRIOR APPLICATION NUMBER: US 09/127,444
; PRIOR FILING DATE: 1998-08-21
; NUMBER OF SEQ ID NOS: 72
; SEQ ID NO 4
; LENGTH: 144
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; NAME/KEY: DOMAIN
; OTHER INFORMATION: Amino acid sequence of stratum corneum chymotryptic
; OTHER INFORMATION: enzyme (scce) catalytic domain
US-10-461-787-4

Query Match          100.0%; Score 40; DB 14; Length 144;
Best Local Similarity 100.0%; Pred. No. 0.97;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 RLSSMVKKV 9
Db      60 RLSSMVKKV 68

RESULT 5
US-09-888-615-98
; Sequence 98, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: FLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION:
US-09-888-615-98

Query Match          100.0%; Score 40; DB 9; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches          9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 RLSSMVKKV 9
Db      123 RLSSMVKKV 131

RESULT 6
US-09-764-762-3
; Sequence 3, Application US/09764762
; Patent No. US20020068341A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NOVEL KALLIKREIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
```

```

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/764,762
; FILING DATE: 16-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/210,084
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PP-0252 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 532504
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-764-762-3

Query Match 100.0%; Score 40; DB 9; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSWVKV 9
DB 123 RLSSWVKV 131

RESULT 7
US-10-264-283-90
; Sequence 90, Application US/10264283
; Publication No. US20030144494A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.590
; CURRENT APPLICATION NUMBER: US/10/264,283
; CURRENT FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 90
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-283-90

Query Match 100.0%; Score 40; DB 14; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSWVKV 9
DB 123 RLSSWVKV 131

RESULT 8
US-10-295-027-498
; Sequence 498, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natacha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 498
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-498

Query Match 100.0%; Score 40; DB 15; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSWVKV 9
DB 123 RLSSWVKV 131

RESULT 9
US-10-173-999-48
; Sequence 48, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002420US
; CURRENT APPLICATION NUMBER: US/10/173,999
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13

```

;; PRIOR APPLICATION NUMBER: US 60/372,246
;; PRIOR FILING DATE: 2001-04-12
;; NUMBER OF SEQ ID NOS: 163
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 48
;; LENGTH: 253
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-173-999-48

Query Match 100.0%; Score 40; DB 15; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.7; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

QY 1 RLSSMVKKV 9
DB 123 RLSSMVKKV 131

RESULT 10
US-09-918-243-110
; Sequence 110, Application US/09918243
; Publication No. US20020142317A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 110
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 122-130 of the SCCE protein
US-09-918-243-110

Query Match 90.0%; Score 36; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKK 8
DB 2 RLSSMVKK 9

RESULT 11
US-09-905-083-110
; Sequence 110, Application US/09905083
; Patent No. US20020146708A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP/C/Div
; CURRENT APPLICATION NUMBER: US/09/905,083
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 09/502,600
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 110
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 122-130 of the SCCE protein

US-09-905-083-110

Query Match 90.0%; Score 36; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKK 8
DB 2 RLSSMVKK 9

RESULT 12
US-10-094-240-4
; Sequence 4, Application US/10094240
; Publication No. US20030082637A1
; GENERAL INFORMATION:
; APPLICANT: ZWIEBEL, LAURENCE J.
; TITLE OF INVENTION: ARRESTIN GENE, POLYPEPTIDE, AND METHODS OF USE THEREOF
; FILE REFERENCE: N8289
; CURRENT APPLICATION NUMBER: US/10/094,240
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 10/056,405
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: 60/264,649
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Anopheles gambiae
US-10-094-240-4

Query Match 82.5%; Score 33; DB 14; Length 394;
Best Local Similarity 77.8%; Pred. No. 80;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9
DB 207 RLSSMVKKL 215

RESULT 13
US-10-056-405-4
; Sequence 4, Application US/10056405
; Publication No. US20030166013A1
; GENERAL INFORMATION:
; APPLICANT: ZWIEBEL, LAURENCE J.
; TITLE OF INVENTION: MOSQUITO OLFACTORY GENES, POLYPEPTIDES, AND METHODS OF
; FILE REFERENCE: N7841
; CURRENT APPLICATION NUMBER: US/10/056,405
; CURRENT FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: 60/264,649
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Anopheles gambiae
US-10-056-405-4

Query Match 82.5%; Score 33; DB 14; Length 394;
Best Local Similarity 77.8%; Pred. No. 80;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9
DB 207 RLSSMVKKL 215

RESULT 14

US-09-918-243-79
; Sequence 79; Application US/09918243
; Patent No. US20020142317A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 79
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 125-133 of the SCCE protein
US-09-918-243-79

Query Match 77.5%; Score 31; DB 9; Length 9;
Best Local Similarity 100.0%; Pred.No. 7.1e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SSMVKKV 9
Db 1 SSMVKKV 7

RESULT 15
US-09-918-243-87
; Sequence 87; Application US/09918243
; Patent No. US20020142317A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 87
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 125-133 of the SCCE protein
US-09-918-243-87

Query Match 77.5%; Score 31; DB 9; Length 9;
Best Local Similarity 100.0%; Pred.No. 7.1e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SSMVKKV 9
Db 1 SSMVKKV 7

Search completed: March 1, 2004, 18:08:49
Job time : 25.1111 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:21:01 ; Search time 10.2222 Seconds
(without alignments)
84.690 Million cell updates/sec

Title: US-09-905-083-32

Perfect score: 40

Sequence: 1 RLSSMVKKV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78.*

1: Pirl.*

2: Pirl2.*

3: Pirl3.*

4: Pirl4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	253	2 A53968	serine proteinase
2	33	82.5	282	2 T35294	probable endo alph
3	33	82.5	372	2 S20036	para-hydroxybenzoa
4	33	82.5	560	2 S50401	hypothetical prote
5	32	80.0	350	1 C70009	ABC transporter (1
6	32	80.0	351	2 D90264	biotin synthase (b
7	32	80.0	358	2 T38914	para-hydroxybenzoa
8	32	80.0	589	2 A71277	arginine-tRNA liga
9	32	80.0	664	2 T50316	hypothetical Armad
10	31	77.5	912	2 H85470	hypothetical prote
11	31	77.5	925	2 T05012	hypothetical prote
12	30	75.0	239	2 T02473	hypothetical prote
13	30	75.0	240	2 T32363	hypothetical prote
14	30	75.0	318	2 T56519	taipoxin-associate
15	30	75.0	319	2 JC5402	vitamin D receptor
16	30	75.0	373	2 C75073	hypothetical prote
17	30	75.0	457	2 D85429	cytochrome P450 11
18	30	75.0	532	2 T01759	glycine hydroxymet
19	30	75.0	647	2 A49218	hemagglutinin homo
20	30	75.0	833	2 A47528	transcription fact
21	30	75.0	986	2 E90220	isoleucine-tRNA sv
22	30	75.0	1167	2 A82543	chromosome segrega
23	30	75.0	1846	2 T33079	hypothetical prote
24	29	72.5	253	2 JC5511	TATA-binding prote
25	29	72.5	285	2 A97255	S-adenosylmethioni
26	29	72.5	336	2 B69214	grp-binding protei
27	29	72.5	337	2 AH2591	membrane lipoprote
28	29	72.5	337	2 E97374	Deinococcus radiod
29	29	72.5	345	2 E90194	threonine synthase

30 29 72.5 349 2 F84246
31 29 72.5 357 2 D96986
32 29 72.5 379 2 H75318
33 29 72.5 553 2 A72228
34 29 72.5 571 2 S65060
35 29 72.5 578 2 H82872
36 29 72.5 582 2 S29314
37 29 72.5 583 2 A45381
38 29 72.5 656 2 T33338
39 29 72.5 691 2 E70906
40 29 72.5 839 2 S62963
41 29 72.5 1360 2 F96596
42 29 72.5 4077 2 T17484
43 29 72.5 4589 2 T14914
44 28 70.0 41 2 G71375
45 28 70.0 63 2 AH2565

hypothetical prote
probable lipoprote
membrane lipoprote
hypothetical prote
phytoene desaturas
hypothetical prote
phytoene dehydroge
phytoene dehydroge
hypothetical prote
probable beta-gluc
hypothetical prote
hypothetical prote
hypothetical prote
dynein beta heavy
hypothetical prote
hypothetical prote

ALIGNMENTS

RESULT 1

A53968

serine proteinase SCCE precursor - human

N;Alternate names: stratum corneum chymotryptic enzyme

C;Species: Homo sapiens (man)

C;Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 22-Jun-1999

C;Accession: A53968

R;Hansson, L.; Stroemqvist, M.; Baackman, A.; Wallbrandt, P.; Carlstein, A.; Egelrud, T.

J. Biol. Chem. 269, 19420-19426, 1994

A;Title: Cloning, expression, and characterization of stratum corneum chymotryptic enzyme

A;Reference number: A53968; MUID:94309225; PMID:8034709

A;Accession: A53968

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-253 <HAN>

A;Cross-references: GB:L33404; NID:G521214; PIDN:AAC37551.1; PID:G532504

C;Genetics:

A;Gene: GDB:PRRS6; SCCE

A;Cross-references: GDB:377730

A;Map position: 7q35-7q35

C;Superfamily: trypsin; trypsin homology

F;30-245/Domain: trypsin homology <TRY>

Query Match 100.0%; Score 40; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9

|||||

Db 123 RLSSMVKKV 131

RESULT 2

T35294

probable endo alpha-1,4 polygalactosaminidase - Streptomyces coelicolor

C;Species: Streptomyces coelicolor

C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999

C;Accession: T35294

R;Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, July 1999

A;Reference number: Z21574

A;Accession: T35294

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-282 <SEE>

A;Cross-references: EMBL:AL096872; PIDN:CAB51262.1; GSPDB:GN00070; SCOREDB:SCSF7.23C

A;Experimental source: strain A3(2)

C;Genetics:

A;Gene: SCOREDB:SCSF7.23C

Query Match 82.5%; Score 33; DB 2; Length 282;
Best Local Similarity 87.5%; Pred. No. 13;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMWKKV 8
|:|||||
Db 265 RLSSMLKK 272

RESULT 3
S20056
Para-hydroxybenzoate polyphenyltransferase (EC 2.5.1.1-) precursor - yeast (Saccharomyces cerevisiae)
N;Alternate names: 4-hydroxybenzoate hexaprenyltransferase; protein YNR0419; protein YNR0419
C;Species: Saccharomyces cerevisiae
C;Date: 23-Apr-1993 #sequence_revision 23-Apr-1993 #text_change 21-Jul-2000
C;Accession: S20056; S63372; J02317
R;Ashby, M.N.; Kutsunai, S.Y.; Ackerman, S.; Tzagoloff, A.; Edwards, P.A.
J. Biol. Chem. 267, 4128-4136, 1992
A;Title: COO2 is a candidate for the structural gene encoding para-hydroxybenzoate:polyP
A;Reference number: S20056; MUID:92156158; PMID:1740455
A;Accession: S20056
A;Molecule type: DNA
A;Residues: 1-372 <ASH>
A;Cross-references: EMBL:M81698; NID:g171252; PIDN:AAA34507.1; PID:g171254
R;Pohl, T.M.
submitted to the Protein Sequence Database, April 1996
A;Reference number: S63346
A;Accession: S63372
A;Molecule type: DNA
A;Residues: 1-372 <POH>
A;Cross-references: EMBL:D71656; NID:g1302546; PID:e239832; PID:g1302547; MIPS:YNR041c
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:COQ2
A;Cross-references: SGD:S0005324; MIPS:YNR041c
A;Map position: 14R
C;Keywords: isoprenoid biosynthesis; mitochondrion; transferase; transmembrane protein
F;1-22/Domain: transit peptide (mitochondrion) #status predicted <TRP>
F;117-133/Domain: transmembrane #status predicted <TM1>
F;172-188/Domain: transmembrane #status predicted <TM2>
F;203-225/Region: polyphenyl diphosphate binding #status predicted

Query Match 82.5%; Score 33; DB 2; Length 372;
Best Local Similarity 77.8%; Pred. No. 17;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMWKKV 9
|:|||||
Db 326 RLFSMLKK 334

RESULT 4
S50401
hypoetical protein YMR145c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein YMR145c
C;Species: Saccharomyces cerevisiae
C;Date: 13-Jan-1995 #sequence_revision 24-Feb-1995 #text_change 19-Apr-2002
C;Accession: S50401
R;Radcock, K.; Churcher, C.
submitted to the EMBL Data Library, December 1994
A;Reference number: S50388
A;Accession: S50401
A;Molecule type: DNA
A;Residues: 1-560 <BAD>
A;Cross-references: EMBL:Z47071; NID:g606429; PIDN:CAA87359.1; PID:g606443; GSPDB:GN0001
C;Genetics:
A;Gene: MIPS:YMR145C
A;Cross-references: SGD:S0004753
A;Map position: 13R
C;Superfamily: NADH dehydrogenase

Query Match 82.5%; Score 33; DB 2; Length 560;
Best Local Similarity 77.8%; Pred. No. 25;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMWKKV 9
|:|||||
Db 347 RLKTWKKV 355

RESULT 5
C70009
ABC transporter (lipoprotein) homolog ynfN - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C;Accession: C70009
R;Kunat, F.; Ogasawara, N.; Mosser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillette, S.; Bruschi, C.V.; Caldwell, B.; Capuani, V.; Carter, N.M.; Choi
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Gallero
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.;
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel,
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle,
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seror,
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpsstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A65580; MUID:98044033; PMID:9384377
A;Accession: C70009
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-350 <XUN>
A;Cross-references: GB:Z99120; GB:AL009126; NID:g2635613; PIDN:CAB15143.1; PID:g2635650
A;Experimental source: strain 168
C;Genetics:
A;Gene: ynfN
C;Superfamily: ABC transporter ynfN

Query Match 80.0%; Score 32; DB 1; Length 350;
Best Local Similarity 87.5%; Pred. No. 27;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSSMWKKV 9
|:|||||
Db 269 LSSMWKKV 276

RESULT 6
D90264
bionin synthase (bioB) [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C;Accession: D90264
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-v
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: D90264
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-351 <KUR>
A;Cross-references: GB:AE006641; NID:g13814305; PIDN:AAK41371.1; GSPDB:GN00155
C;Genetics:
A;Gene: bioB

Query Match 80.0%; Score 32; DB 2; Length 351;
Best Local Similarity 66.7%; Pred. No. 27;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMWKKV 9
|:|||||
Db 341 RLDSMLKKV 349

RESULT 7
T38914
Para-hydroxybenzoate-polyphenyltransferase (EC 2.5.1.1) precursor, mitochondrial - fission
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T38914
R:Pearson, D.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, February 1996
A:Reference number: 221817
A:Accession: T38914
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-358 <PEA>
A:Cross-references: EMBL:Z69728; PIDN:CAA93575.1; GSPDB:GN00066; SPDB:SPAC56F8.04C
A:Experimental source: strain 972h-; cosmid c56F8
C:Genetics:
A:Gene: SPDB:SPAC56F8.04C
A:Map position: 1
A:Genome: nuclear
C:Keywords: mitochondrion; transferase

Query Match 80.0%; Score 32; DB 2; Length 358;
Best Local Similarity 77.8%; Pred. No. 28;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMWKVKV 9
DB 311 RLSSMIYKVK 319

RESULT 8
A71277
arginine-tRNA ligase (EC 6.1.1.19) (args) - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 03-Jun-2002
C:Accession: A71277
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khatak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770; PMID:9665876
A:Accession: A71277
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-589 <COL>
A:Cross-references: GB:AE001253; GB:AE000520; NID:g3323133; PIDN:AAC65797.1; PID:g332314
A:Experimental source: strain Nichols
C:Genetics:
A:Gene: TP0831
C:Superfamily: Bacillus arginine-tRNA ligase
C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 80.0%; Score 32; DB 2; Length 589;
Best Local Similarity 66.7%; Pred. No. 45;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMWKVKV 9
DB 473 RLSSLLKVK 481

RESULT 9
T50316
hypothetical Armadillo/beta-catenin domain protein [imported] - fission yeast (Schizosac
C:Species: Schizosaccharomyces pombe
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jun-2000
C:Accession: T50316
R:McGough, R.C.; Rajandream, M.A.; Barrell, B.G.; Cadieu, E.; Lelaure, V.; Galibert, F
submitted to the EMBL Data Library, January 2000
A:Reference number: 225061

Query Match 77.5%; Score 31; DB 2; Length 925;
Best Local Similarity 55.6%; Pred. No. 1.2e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMWKVKV 9
DB 403 RLSSMIRKI 411

RESULT 11
T05012
hypothetical protein T19P19.140 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C:Accession: T05012
R:Bevan, M.; Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Hoheisel, J.; Mewes, H.W.; J
submitted to the Protein Sequence Database, April 1998
A:Reference number: Z15394
A:Accession: T05012
A:Molecule type: DNA
A:Residues: 1-925 <BEV>
A:Cross-references: EMBL:AL022605
A:Experimental source: cultivar Columbia; BAC clone T19P19
C:Genetics:
A:Map position: 4
A:Introns: 142/3; 193/1; 551/1
A>Note: T19P19.140

Query Match 77.5%; Score 31; DB 2; Length 925;
Best Local Similarity 55.6%; Pred. No. 1.2e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMWKVKV 9
DB 403 RLSSMIRKI 411

RESULT 10
H85470
hypothetical protein ATg39750 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: H85470
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Harbor, NY, USA
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: H85470
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-912 <STO>
A:Cross-references: GB:NC_001268; NID:g7270959; PIDN:CAB80638.1; GSPDB:GN00140
C:Genetics:
A:Gene: ATg39750
A:Map position: 4

Query Match 77.5%; Score 31; DB 2; Length 912;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMWKVKV 9
DB 403 RLSSMIRKI 411

RESULT 11
T05012
hypothetical protein T19P19.140 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C:Accession: T05012
R:Bevan, M.; Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Hoheisel, J.; Mewes, H.W.; J
submitted to the Protein Sequence Database, April 1998
A:Reference number: Z15394
A:Accession: T05012
A:Molecule type: DNA
A:Residues: 1-925 <BEV>
A:Cross-references: EMBL:AL022605
A:Experimental source: cultivar Columbia; BAC clone T19P19
C:Genetics:
A:Map position: 4
A:Introns: 142/3; 193/1; 551/1
A>Note: T19P19.140

Query Match 77.5%; Score 31; DB 2; Length 925;
Best Local Similarity 55.6%; Pred. No. 1.2e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMWKVKV 9
DB 403 RLSSMIRKI 411

RESULT 10
H85470
hypothetical protein ATg39750 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: H85470
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Harbor, NY, USA
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: H85470
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-912 <STO>
A:Cross-references: GB:NC_001268; NID:g7270959; PIDN:CAB80638.1; GSPDB:GN00140
C:Genetics:
A:Gene: ATg39750
A:Map position: 4

Query Match 77.5%; Score 31; DB 2; Length 912;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMWKVKV 9
DB 403 RLSSMIRKI 411

RESULT 11
T05012
hypothetical protein T19P19.140 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C:Accession: T05012
R:Bevan, M.; Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Hoheisel, J.; Mewes, H.W.; J
submitted to the Protein Sequence Database, April 1998
A:Reference number: Z15394
A:Accession: T05012
A:Molecule type: DNA
A:Residues: 1-925 <BEV>
A:Cross-references: EMBL:AL022605
A:Experimental source: cultivar Columbia; BAC clone T19P19
C:Genetics:
A:Map position: 4
A:Introns: 142/3; 193/1; 551/1
A>Note: T19P19.140

Query Match 77.5%; Score 31; DB 2; Length 925;
Best Local Similarity 55.6%; Pred. No. 1.2e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMWKVKV 9
DB 403 RLSSMIRKI 411

RESULT 10
H85470
hypothetical protein ATg39750 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: H85470
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Harbor, NY, USA
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: H85470
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-912 <STO>
A:Cross-references: GB:NC_001268; NID:g7270959; PIDN:CAB80638.1; GSPDB:GN00140
C:Genetics:
A:Gene: ATg39750
A:Map position: 4

Query Match 77.5%; Score 31; DB 2; Length 912;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMWKVKV 9
DB 403 RLSSMIRKI 411

RESULT 11
T05012
hypothetical protein T19P19.140 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C:Accession: T05012
R:Bevan,

Db 416 RFSSMIRKI 424

RESULT 12

T02473

hypothetical protein At2g45740 [imported] - Arabidopsis thaliana

N;Alternate names: hypothetical protein F4118.28

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001

C;Accession: T02473; C84894

R;Rounsley, S.D.; Lin, X.; Kechum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul

submitted to the EMBL Data Library, August 1998

A;Description: Arabidopsis thaliana chromosome II BAC F4118 genomic sequence.

A;Reference number: Z14674

A;Accession: T02473

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-239 <ROU>

A;Cross-references: EMBL:AC004665; NID:g3386593; PID:g3386621

A;Experimental source: cultivar Columbia

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: C84894

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-239 <STO>

A;Cross-references: GB:AE002093; NID:g3386621; PIDN:AAC28551.1; GSPDB:GN00139

C;Genetics:

A;Gene: At2g45740; F4118.28

A;Map position: 2

A;Introns: 65/3; 90/3; 115/3; 142/3; 167/3

Query Match 75.0%; Score 30; DB 2; Length 239;

Best Local Similarity 66.7%; Pred. No. 53;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9

Db 148 RLSSMKKI 156

RESULT 13

T32363

hypothetical protein C08E3.12 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C;Accession: T32363

R;Miller, N.; Kramer, J.; Keppler, D.

submitted to the EMBL Data Library, September 1997

A;Description: The sequence of C. elegans cosmid C08E3.

A;Reference number: Z21155

A;Accession: T32363

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-240 <MIL>

A;Cross-references: EMBL:AF025457; PIDN:AAB70970.1; GSPDB:GN00020; CESP:C08E3.12

C;Genetics:

A;Gene: CESP:C08E3.12

A;Map position: 2

A;Introns: 167/3

Query Match 75.0%; Score 30; DB 2; Length 240;

Best Local Similarity 66.7%; Pred. No. 53;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9

Db 193 RIPSMLKKV 201

RESULT 14

T56519

taipoxin-associated calcium binding protein-49 precursor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 02-Aug-2002

C;Accession: T56519

J;Dodds, D.; Schlimgen, A.K.; Lu, S.Y.; Perin, M.S.

R;Neurochem. 64, 2339-2344, 1995

A;Title: Novel reticular calcium binding protein is purified on taipoxin columns.

A;Reference number: T56519; MUID:95239201; PMID:7722520

A;Accession: T56519

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-318 <RES>

A;Cross-references: EMBL:U15734; NID:g606967; PIDN:AAA80197.1; PID:g606968

C;Superfamily: reticulocalbin; calmodulin repeat homology

C;Keywords: calcium binding; EF hand

F;62-94/Domain: calmodulin repeat homology <EF1>

F;128-152/Domain: calmodulin repeat homology <EF2>

F;187-219/Domain: calmodulin repeat homology <EF3>

F;228-260/Domain: calmodulin repeat homology <EF4>

F;262-296/Domain: calmodulin repeat homology <EF5>

Query Match 75.0%; Score 30; DB 2; Length 318;

Best Local Similarity 55.6%; Pred. No. 69;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9

Db 66 RLQSIIRKI 74

RESULT 15

JC5402

vitamin D receptor associated factor 1 - mouse

N;Alternate names: ERC-55 protein

C;Species: Mus musculus (house mouse)

C;Date: 07-Jul-1997 #sequence_revision 18-Jul-1997 #text_change 02-Aug-2002

C;Accession: JC5402

R;Imai, T.; Matsuda, K.; Shimojima, T.; Hashimoto, T.; Masuhiro, Y.; Kitamoto, T.; Sugita

Biochem. Biophys. Res. Commun. 233, 765-769, 1997

A;Title: ERC-55, a binding protein for the papilloma virus E6 oncoprotein, specifically,

A;Reference number: JC5402; MUID:97312489; PMID:9168930

A;Accession: JC5402

A;Molecule type: mRNA

A;Residues: 1-319 <IMA>

C;Comment: This protein acts as a vitamin D receptor-specific cofactor modulating its fur

C;Superfamily: reticulocalbin; calmodulin repeat homology

C;Keywords: EF hand

F;63-95/Domain: calmodulin repeat homology <EF1>

F;99-131/Domain: calmodulin repeat homology <EF2>

F;151-183/Domain: calmodulin repeat homology <EF3>

F;188-220/Domain: calmodulin repeat homology <EF4>

F;229-261/Domain: calmodulin repeat homology <EF5>

F;265-297/Domain: calmodulin repeat homology <EF6>

Query Match 75.0%; Score 30; DB 2; Length 319;

Best Local Similarity 55.6%; Pred. No. 70;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9

Db 67 RLQSIIRKI 75

Search completed: March 1, 2004, 17:36:26

Job time : 12.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:17:25 ; Search time 6 Seconds
(without alignments)
78.105 Million cell updates/sec

Title: US-09-905-083-32
Perfect score: 40
Sequence: 1 RLSSMVKKV 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	40	100.0	253	1 KX77_HUMAN	P49862 homo sapien
2	33	82.5	372	1 COQ2_YEAST	P32378 saccharomyc
3	33	82.5	560	1 YU23_YEAST	P40215 saccharomyc
4	32	80.0	350	1 YUEN_BACSU	O05252 bacillus su
5	32	80.0	358	1 COQ2_SCHPO	O10252 schizosacch
6	32	80.0	589	1 SYR_TREPA	O83803 treponema p
7	30	75.0	318	1 RCN2_RAT	O62703 rattus norv
8	30	75.0	447	1 EX7L_STRMU	O62703 streptococc
9	30	75.0	524	1 C90C_ARATH	O9m066 arabidopsis
10	30	75.0	572	1 SNTD_LUTLO	O9xz43 lutzomyia l
11	30	75.0	833	1 MBP1_YEAST	P39678 saccharomyc
12	29	72.5	253	1 T2D7_RAT	O62880 rattus norv
13	29	72.5	351	1 EGC2_OCEIH	O8etb7 oceanobacil
14	29	72.5	571	1 CRT1_MAIZE	P49086 zea mays (m
15	29	72.5	582	1 CRT1_CAPAN	P80093 capsulicum an
16	29	72.5	583	1 CRT1_LYCES	P28554 lycopersico
17	29	72.5	839	1 COD2_YEAST	P53959 saccharomyc
18	28	70.0	41	1 Y010_TREPA	O83054 treponema p
19	28	70.0	109	1 YES7_YEAST	P40062 saccharomyc
20	28	70.0	156	1 ATPF_BUCBP	O89b43 buchnera ap
21	28	70.0	175	1 PAP1_MOUSE	P35230 mus musculu
22	28	70.0	213	1 MAG2_SCHPO	O94468 schizosacch
23	28	70.0	264	1 RS3A_APLCA	P49395 aplysia cal
24	28	70.0	317	1 FABD_BACSU	P71019 bacillus su
25	28	70.0	334	1 RUVE_STRAAM	O98t12 staphylococ
26	28	70.0	334	1 RUVE_STRAAM	O98t12 staphylococ
27	28	70.0	448	1 FIAP_ADRP3	O83457 porcine ade
28	28	70.0	459	1 TRME_STRAAM	O931e1 staphylococ
29	28	70.0	459	1 TRME_STRAAM	O99q33 staphylococ
30	28	70.0	459	1 TRME_STABPP	O8cmh5 staphylococ
31	28	70.0	473	1 6PGD_BUCAP	O9zhd9 buchnera ap
32	28	70.0	506	1 GTA_NPVAC	P41447 autographa
33	28	70.0	523	1 CXK3_ARATH	O9lts3 arabidopsis

34	28	70.0	524	1 IMD4_YEAST	P50094 saccharomyc
35	28	70.0	553	1 YM8A_YEAST	O04847 saccharomyc
36	28	70.0	583	1 SYR_CANBF	O7vqx7 candidatus
37	28	70.0	625	1 HTPG_BACHD	O9kes1 bacillus ha
38	28	70.0	690	1 PTA_MYCTU	P96254 mycobacteri
39	28	70.0	758	1 K173_HUMAN	O14684 homo sapien
40	28	70.0	851	1 YD72_SCHPO	Q10377 schizosacch
41	28	70.0	1021	1 BUB1_YEAST	P41695 saccharomyc
42	28	70.0	1163	1 LEPR_MACMU	O9my10 macaca mula
43	28	70.0	1165	1 LEPR_HUMAN	P48357 homo sapien
44	27	67.5	60	1 68MP_BOVIN	P14790 bos taurus
45	27	67.5	89	1 RS20_HELPJ	O9zms1 helicobacte

ALIGNMENTS

RESULT 1
KLK7_HUMAN
ID KLK7_HUMAN STANDARD; PRT; 253 AA.
AC P49862;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Kallikrein 7 precursor (EC 3.4.21.-) (Stratum corneum chymotryptic enzyme) (hSCCE).
DE KLK7 OR PRSS6 OR SCCE.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-53.
RC TISSUE=Skin;
RX MEDLINE=94308225; PubMed=8034709;
RA Hansson L., Stroemqvist M., Baeckman A., Wallbrandt P., Carlstein A., Egelrud T.;
RT "Cloning, expression, and characterization of stratum corneum chymotryptic enzyme. A skin-specific human serine proteinase.";
RL J. Biol. Chem. 269:19420-19426(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Keratinocytes;
RA Yousef G.M., Scorilas A., Diamandis E.P.;
RT "Molecular characterization, mapping and tissue expression of the human stratum corneum chymotryptic enzyme gene.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC MEDLINE=20510030; PubMed=11054574;
RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J., Moss P., Paepker B., Wang K.;
RT "Sequencing and expression analysis of the serine protease gene cluster located in chromosome 19q13 region.";
RL Gene 257:119-130(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC MEDLINE=95314630; PubMed=7794273;
RA Hansson L., Baekman A., Ny A., Edlund M., Edholm E., Tornell J., Wallbrandt P., Egelrud T.;
RT "Epidermal overexpression of stratum corneum chymotryptic enzyme in mice; a model for chronic itchy dermatitis.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP CHARACTERIZATION.
RC MEDLINE=95314630; PubMed=7794273;
RA Skytt A., Stroemqvist M., Egelrud T.;
RT "Primary substrate specificity of recombinant human stratum corneum chymotryptic enzyme.";
RL Biochem. Biophys. Res. Commun. 211:586-589(1995).
CC !- FUNCTION: May catalyze the degradation of intercellular cohesive structures in the cornified layer of the skin in the continuous shedding of cells from the skin surface. Specific for amino acid residues with aromatic side chains in the P1 position. SCCE

CC cleaves insulin B chain at 6-Leu-Cys-7, 16-Tyr-Leu-17, 25-Phe-
 CC [-Tyr-26, and 26-Tyr-Thr-27. Could play a role in the activation
 CC of precursors to inflammatory cytokines.
 CC -!- TISSUE SPECIFICITY: Abundantly expressed in the skin and is
 CC expressed by keratinocytes in the epidermis. Very low levels are
 CC also seen in the brain and kidney.
 CC -!- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L33404; AAC37551.1; -;
 CC EMBL; AF166330; RAD49718.1; -;
 CC EMBL; AF243527; AAG33360.1; -;
 CC EMBL; AF332583; AAK69624.1; -;
 CC PIR; A53968; A53968.
 CC HSSP; P00763; LDPO.
 CC MEROPS; S01.300; -;
 CC Genew; HGNC:6368; KLIK7.
 CC MIM; 604438; -;
 CC GO; GO:0008236; F:serine-type peptidase activity; TAS.
 CC GO; GO:0008544; P:epidermal differentiation; TAS.
 CC InterPro; IPR009003; Cys Ser trypsin.
 CC InterPro; IPR001254; Peptidase_S1.
 CC InterPro; IPR001314; Peptidase_S1A.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC SMART; SM00020; TRYP_SPC; 1.
 CC PROSITE; PS00240; TRYPSIN_DOM; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 CC Hydrolase; Serine protease; Zymogen; Glycoprotein; Signal.
 CC SIGNAL 1 22
 CC FT PRINTS; PR00722; CHYMOTRYPSIN.
 CC FT PROPEP 23 29
 CC FT CHAIN 30 253
 CC KALLIKREIN 7.
 CC ACT_SITE 70 70
 CC CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC FT ACT_SITE 112 112
 CC CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC FT ACT_SITE 205 205
 CC CHARGE RELAY SYSTEM (BY SIMILARITY).
 CC FT DISULFID 36 137
 CC BY SIMILARITY.
 CC FT DISULFID 55 171
 CC BY SIMILARITY.
 CC FT DISULFID ? 239
 CC BY SIMILARITY.
 CC FT DISULFID 144 211
 CC BY SIMILARITY.
 CC FT DISULFID 176 190
 CC BY SIMILARITY.
 CC FT DISULFID 201 226
 CC BY SIMILARITY.
 CC FT CARBOHYD 246 246
 CC N-LINKED (GLCNAC...) (POTENTIAL).
 CC SQ SEQUENCE 253 AA; 27525 MW; 2D68B6B15A76A668 CRC64;
 CC
 CC Query Match 100.0%; Score 40; DB 1; Length 253;
 CC Best Local Similarity 100.0%; Pred. No. 0.2; Indels 0; Gaps 0;
 CC Matches 9; Conservative 0; Mismatches 0;
 CC
 CC QY 1 RLSSMWKXV 9
 CC
 CC Db 123 RLSSMWKXV 131
 CC
 CC RESULT 2
 CC COQ2 YEAST
 CC ID COQ2 YEAST STANDARD; PRT; 372 AA.
 CC AC P32378;
 CC DT 01-OCT-1993 (Rel. 27, Created)
 CC DT 01-OCT-1993 (Rel. 27, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Para-hydroxybenzoate--polyprenyltransferase, mitochondrial precursor
 CC (EC 2.5.1.1) (PDB:polyprenyltransferase).
 CC DE COQ2 OR YNR041C OR N3419.
 CC GN Saccharomycetes cerevisiae (Baker's yeast).
 CC OS Saccharomycetes; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Eukaryota;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN
 RN SEQUENCE FROM N.A.
 RX MEDLINE92156158; PubMed1740455;
 RA Ashby M.N., Kutsunai S.Y., Ackerman S., Tzagoloff A., Edwards P.A.;
 RT "COQ2 is a candidate for the structural gene encoding para-
 RT hydroxybenzoate:polyprenyltransferase.";
 RL J. Biol. Chem. 267:4128-4136(1992).
 RN [2]
 RN SEQUENCE FROM N.A.
 RA Pohl T.M.;
 RP Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CATALYZES THE PRENYLATION OF PARA-HYDROXYBENZOATE
 CC WITH AN ALL-TRANS POLYPRENYL GROUP.
 CC -!- PATHWAY: Coenzyme Q biosynthesis; second step.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane.
 CC -!- SIMILARITY: Belongs to the ubiA prenyltransferase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M81698; AAA34507.1; -;
 CC EMBL; Z71656; CAA96321.1; -;
 CC PIR; S20056; S20056.
 CC GeneOnline; 143386; -;
 CC SGD; S0005324; COQ2.
 CC GO; GO:0004659; F:prenyltransferase activity; IMP.
 CC GO; GO:0006743; P:ubiquinone metabolism; IMP.
 CC InterPro; IPR000537; UbiA.
 CC InterPro; IPR006370; UbiA_proteo.
 CC Pfam; PF01040; UbiA; 1.
 CC TIGRFAMs; TIGR01474; ubiA_proteo; 1.
 CC PROSITE; PS00943; UbiA; 1.
 CC Isoprene biosynthesis; Transferase; Transf peptide; Mitochondrion;
 CC Transmembrane.
 CC FT TRANSIT 1 22 MITOCHONDRION (POTENTIAL).
 CC FT CHAIN 23 372 PARA-HYDROXYBENZOATE--
 CC POLYPRENYLTRANSFERASE.
 CC FT TRANSMEM 92 112 1 (POTENTIAL).
 CC FT TRANSMEM 114 133 2 (POTENTIAL).
 CC FT TRANSMEM 170 191 3 (POTENTIAL).
 CC FT TRANSMEM 193 210 4 (POTENTIAL).
 CC FT TRANSMEM 229 249 5 (POTENTIAL).
 CC FT TRANSMEM 298 318 6 (POTENTIAL).
 CC FT DOMAIN 29 49 SER-RICH.
 CC FT DOMAIN 134 156 ALLYLIC POLYPRENYL DIPHOSPHATE-BINDING
 CC SITE (POTENTIAL).
 CC SQ SEQUENCE 372 AA; 41001 MW; 4D9738CE248B4AD6 CRC64;
 CC
 CC Query Match 82.5%; Score 33; DB 1; Length 372;
 CC Best Local Similarity 77.8%; Pred. No. 9.9; Indels 0; Gaps 0;
 CC Matches 7; Conservative 1; Mismatches 1;
 CC
 CC QY 1 RLSSMWKXV 9
 CC
 CC Db 326 RLFSNIKKV 334
 CC
 CC RESULT 3
 CC YM23 YEAST
 CC ID YM23 YEAST STANDARD; PRT; 560 AA.
 CC AC P40315;
 CC DT 01-FEB-1995 (Rel. 31, Created)
 CC DT 01-FEB-1995 (Rel. 31, Last sequence update)
 CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
 CC DE Hypothetical 62.8 kDa protein in RPS16A-TIF34 intergenic region.
 CC DE

GN YMR145C OR YMR375.14C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C / AS972;
RX MEDLINE=97313268; PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagers K., Lye G., Moule S., Odell C., Pearson D., Rajadream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT XIII.";
RL Nature 387:90-93 (1997).
CC -!- SIMILARITY: STRONG, TO YEAST ROTENONE-INSENSITIVE NADH-UBIQUINONE
CC OXIDOREDUCTASE (ND11).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z47071; CAA87359.1; -;
CC FIR; S50401; S50401.
CC GeneOnline; 142815; -;
CC SGD; S0004753; YMR145C.
CC GO; GO:0005739; C:mitochondrion; IDA.
CC GO; GO:0003954; F:NADH dehydrogenase activity; IDA.
CC GO; GO:0019655; P:ethanol fermentation; IMP.
CC GO; GO:0006116; P:NADH oxidation; IDA.
CC InterPro; IPR001327; FAD_Py_redox.
CC Pfam; PF007070; py_redox; 1.
CC Hypothetical protein; Oxidoreductase; Flavoprotein; FAD; NAD;
KW UBIQUINONE.
SQ SEQUENCE 560 AA; 62774 MW; 10B1795E12E29C34 CRC64;

Query Match 82.5%; Score 33; DB 1; Length 560;
Best Local Similarity 77.8%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9
DB 347 RLKTMVKV 355

RESULT 4
YUFG_BACSU STANDARD; PRT; 350 AA.
AC 005252;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical lipoprotein yufN precursor.
GN YUFG OR BSU31540.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=97419515; PubMed=9274030;
RA Oudega B., Koningsstein G., Rodrigues L., de Sales Ramon M.,
RA Hilbert H., Duesterhoef A., Pohl T.M., Weitzenecker T.;
RA "Analysis of the Bacillus subtilis genome: cloning and nucleotide
RT sequence of a 62 kb region between 275 degrees (rrnB) and 284 degrees
RT (pai).";
RL Microbiology 143:2769-2774 (1997).
RN [2]

RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Mosser I., Albertini A.M., Alloni G.,
RA Azevedo V., Bextero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriass R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Bouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Deniot F., Devine K.M., Dusterhoef A., Ehrlich S.D., Emmerson P.T.,
RA Estlin K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
RA Gim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karanata D., Kasahara Y., Klaer-Blanchard M., Klein C.,
RA Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeder R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin S.S., Soldo B.,
RA Sorokin M., Taconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RL Nature 390:249-256 (1997).
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Potential).
CC -!- SIMILARITY: Belongs to the BMP lipoprotein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z93937; CAB07936.1; -;
CC EMBL; Z99120; CAB15143.1; -;
CC PIR; C70009; C70009.
CC Subtilist; BG12349; yufN.
CC InterPro; IPR003760; Bmp.
CC InterPro; IPR000437; Prok_lipoprot_S.
CC Pfam; PF02808; Bmp; 1.
CC PROSITE; PS00013; PROKAR LIPOPROTEIN; FALSE NEG.
CC KW Hypothetical protein; Membrane; Lipoprotein; Signal;
CC Complete proteome; Palmitate.
FT SIGNAL 1 13 PROBABLE.
FT CHAIN 14 350 HYPOTHETICAL LIPOPROTEIN YUFG.
FT LIPID 14 14 N-palmitoyl cysteine (Probable).
FT LIPID 14 14 S-diacylglycerol cysteine (Probable).
SQ SEQUENCE 350 AA; 37349 MW; 16D5176A52A99284 CRC64;

Query Match 80.0%; Score 32; DB 1; Length 350;
Best Local Similarity 87.5%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSSMVKKV 9
DB 269 LTSVMKKV 276

RESULT 5
COQ2_SCHPO

ID AC COO2 SCHPO STANDARD; PRT; 358 AA.
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Para-hydroxybenzoate-polyphenyltransferase, mitochondrial precursor
 DE (EC 2.5.1.1-) (PHE:polyphenyltransferase) (p-hydroxybenzoate polyphenyl
 DE diphosphate transferase).
 GN COO2 OR PPT1 OR SPAC56F8.04C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_taxid=4896;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RX MEDLINE=20545437; PubMed=11092853;
 RA Uchida N., Suzuki K., Saiki R., Kainou T., Tanaka K., Matsuda H.,
 RA Kawamukai M.,
 RT "Phenotypes of fission yeast defective in ubiquinone production due to
 RT disruption of the gene for p-hydroxybenzoate polyphenyl diphosphate
 RT transferase.";
 RL J. Bacteriol. 182:6933-6939(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA McInnes P., Moulis S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
 RA Rutherford K., Rutter S., Saunders R., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymoprez B.,
 RA Weljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabell C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Coffeau A., Cadieu E., Dreano S., Gloux S., Leleau V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
 RA "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -1- FUNCTION: CATALYZES THE PRENYLATION OF PARA-HYDROXYBENZOATE
 CC WITH AN ALL-TRANS POLYPHENYL GROUP (BY SIMILARITY).
 CC -1- PATHWAY: Coenzyme Q biosynthesis; second step.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane.
 CC -1- SIMILARITY: Belongs to the ubia prenyltransferase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB053168; BAB20425.1; -.
 CC FIRM; Z69728; CAA93575.1; -.
 CC FIRM; T38914; T38914.
 CC GenBank; SPAC56F8.04C; -.
 CC InterPro; IPR000537; Ubia.

DR InterPro; IPR006370; Ubia_proteo.
 DR Pfam; PF01040; Ubia; 1.
 DR TIGR; TIGR01474; ubia_proteo; 1.
 DR PROSITE; PS00943; Ubia; 1.
 KW Isoprene biosynthesis; Transferase; Transit peptide; Mitochondrion;
 KW Transmembrane.
 FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
 FT CHAIN ? 358 PARA-HYDROXYBENZOATE--
 FT POLYPHENYLTRANSFERASE.
 FT TRANSMEM 76 96 POTENTIAL.
 FT TRANSMEM 107 127 POTENTIAL.
 FT TRANSMEM 154 174 POTENTIAL.
 FT TRANSMEM 202 222 POTENTIAL.
 FT TRANSMEM 229 249 POTENTIAL.
 FT TRANSMEM 275 295 POTENTIAL.
 FT TRANSMEM 316 336 POTENTIAL.
 FT TRANSMEM 358 358 POTENTIAL.
 SQ SEQUENCE 358 AA; 39454 MW; AA5485411A0922F7 CRC64;
 Query Match 80.0%; Score 32; DB 1; Length 358;
 Best Local Similarity 77.8%; Pred. No. 16;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 RUSSMVKV 9
 DB 311 RUSSMVKV 319
 RESULT 6
 SYR_TREPA STANDARD; PRT; 589 AA.
 ID OC3803;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Arginyl-tRNA synthetase (EC 6.1.1.19) (Arginine-tRNA ligase) (ArgRS).
 DE ARG5 OR TP0831.
 GN Treponema pallidum.
 OS Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
 OX NCBI_taxid=160;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nichols;
 RX MEDLINE=98332770; PubMed=9665876;
 RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
 RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
 RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
 RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,
 RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
 RA Venter J.C.;
 RA "Complete genome sequence of Treponema pallidum, the syphilis
 RA spirochete";
 RL Science 281:375-388(1998).
 CC -1- CATALYTIC ACTIVITY: ATP + L-arginine + tRNA(Arg) = AMP +
 CC diphosphate + L-arginyl-tRNA(Arg).
 CC -1- SUBUNIT: Monomer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB001253; AAC65797.1; -.
 CC FIRM; A71277; A71277.
 CC FIRM; TP0831; -.
 CC HAMAP; MF_00123; -; 1.
 CC InterPro; IPR001278; Arg_tRNA-synt_1c.
 CC InterPro; IPR005148; N.

```
DR InterPro; IPR008909; tRNA-synt_1d_C.
DR InterPro; IPR001412; tRNA-synt_1.
DR Pfam; PF03485; N-Arg; 1.
DR Pfam; PF00750; tRNA-synt_id; 1.
DR Pfam; PF05745; tRNA-synt_id_C; 1.
DR PRINTS; PR01038; TRNASYNTHARG.
DR TIGRFAMs; TIGR00456; argS; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 132 142 "HIGH" REGION.
SQ SEQUENCE 589 AA; 67131 MW; FA6156A69F4568E3 CRC64;

Query Match 80.0%; Score 32; DB 1; Length 589;
Best Local Similarity 66.7%; Pred. NO. 25;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9
Db 473 RLSSLLKKV 481

RESULT 7
RCN2_RAT STANDARD; PRT; 318 AA.
AC Q62703;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Reticulocalbin 2 precursor (Calcium-binding protein ERC-55) (Taipoxin-
DE associated calcium-binding protein-49) (TCBP-49).
GN RCN2 OR ERC55.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=95239201; PubMed=7722520;
RA Dodds D., Schlimgen A.K., Lu S.Y., Perin M.S.;
RT "Novel reticular calcium binding protein is purified on taipoxin
RT columns.";
RL J. Neurochem. 64:2339-2344 (1995).
CC -1- FUNCTION: Not known. Binds calcium.
CC -1- SUBUNIT: Binds taipoxin.
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- SIMILARITY: Belongs to the CREC family.
CC -1- SIMILARITY: Contains 6 EF-hand calcium-binding domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U15734; AAA80137.1; -.
CC FIR; I56519; I56519.
CC InterPro; IPR002048; EF-hand.
CC InterPro; IPR000886; ER_target_S.
CC Pfam; PF00036; ehand; 5.
CC SMART; SM00054; EPH; 3.
CC PROSITE; PS00018; EF_HAND; 5.
CC PROSITE; PS00014; EF_TARGET; 1.
CC Calcium-binding; Endoplasmic reticulum; Signal; Repeat.
KW SIGNAL
FT CHAIN 1 23 POTENTIAL.
FT CHAIN 24 318 RETICULOCALBIN 2.
FT CA_BIND 75 86 EF-HAND 1 (POTENTIAL).
FT CA_BIND 111 122 EF-HAND 2 (POTENTIAL).
FT CA_BIND 163 174 EF-HAND 3 (POSSIBLY ANCESTRAL).
```

```
FT CA_BIND 200 211 EF-HAND 4 (POTENTIAL).
FT CA_BIND 241 252 EF-HAND 5 (POTENTIAL).
FT CA_BIND 277 288 EF-HAND 6 (POTENTIAL).
FT SITE 315 318 PREVENT SECRETION FROM ER (POTENTIAL).
SQ SEQUENCE 318 AA; 37176 MW; 57B50F45FC09CFF CRC64;

Query Match 75.0%; Score 30; DB 1; Length 318;
Best Local Similarity 55.6%; Pred. NO. 39;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9
Db 66 RLQSIKKI 74

RESULT 8
EX7L_STRMU STANDARD; PRT; 447 AA.
ID EX7L_STRMU
AC Q8DV85;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
DE (Exonuclease VII large subunit).
GN XSEA OR SMU.580.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22235063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen.";
RL Proc Natl. Acad. Sci. U.S.A. 99:14434-14439 (2002).
CC -1- FUNCTION: Bidirectionally degrades single-stranded DNA into large
CC acid-insoluble oligonucleotides, which are then degraded further
CC into small acid-soluble oligonucleotides (By similarity).
CC -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-
CC or 3'- to 5'-direction to yield nucleoside 5'-phosphates.
CC -1- SUBUNIT: Heterooligomer composed of large and small subunits (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the xsea family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE014902; AAN58319.1; -.
CC HAMAP; MF 00378; -.
CC InterPro; IPR000674; Aldkan_dh_hamm.
CC InterPro; IPR003753; Exonuc_VII_L.
CC InterPro; IPR008994; Nucleic_acid_OB.
CC InterPro; IPR004365; tRNA anti.
CC Pfam; PF02601; Exonuc_VII_L; 1.
CC Pfam; PF01336; tRNA anti; 1.
CC TIGRFAMs; TIGR00237; xsea; 1.
CC Hydrolase; Nuclease; Exonuclease; Complete proteome.
SQ SEQUENCE 447 AA; 51146 MW; 560412CE7DAFA6A3 CRC64;

Query Match 75.0%; Score 30; DB 1; Length 447;
Best Local Similarity 66.7%; Pred. NO. 54;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```


QY 1 RLSSMVKKV 9
 DB 130 RLSQFVKZI 138

RESULT 9
 C90C_ARATH STANDARD; PRT; 524 AA.

ID C90C_ARATH STANDARD; PRT; 524 AA.

AC Q9066; 023242;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cytochrome P450 90C1 (EC 1.14.-.-) (ROTUNDIFOLIA3).
 GN ROT3 OR CYP90C1 OR ATG36380 OR C7A10.980 OR AP22.10 OR P23B13.220.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98361880; PubMed=9694802;
 RA Kim G.-T., Tsukaya H., Uchimiya H.,
 RA "The ROTUNDIFOLIA3 Gene of Arabidopsis thaliana encodes a new member
 of the cytochrome P-450 family that is required for the regulated
 polar elongation of leaf cells."
 RL Genes Dev. 12:2381-2391 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083489; PubMed=10617198;
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
 RA Poll T., Duetschoff A., Stiekema W., Entian K.-D., Terry N.,
 RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,
 RA Weichselgartner M., de Simone V., Oberwarter B., Mache R., Mueller M.,
 RA Kreis M., Delsen M., Puigdomenech P., Watson M., Schmidtheini T.,
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Kohseil J., Zimmermann W., Wedler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
 RA Van der Schueren J., Crymornez B., Chuang Y.-J., Vandenbusche F.,
 RA Braeken M., Weljens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Weitzneger T., Sothe G., Ramsperger U., Hilbert H., Braun M.,
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirks W.,
 RA Moolenaar P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
 RA Bernier S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 RA De Keyser A., Buyschaert C., Giesen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIlroy K., Mayes R.,
 RA Pettett A., Rajandream M.A., Lyne M., Benes V., Reckmann S.,
 RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
 RA Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
 RA Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,
 RA Schrabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Chedid F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacón D., Jesse T.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielek C.,
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Shest P., Cordes M., Abu-Threideh J.,
 RA Stoneking T., Kalicki J., Graves T., Harnon G., Edwards J.,
 RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Mix P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.W.,
 RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
 RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 RA Granat S., Shohay N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,

Chen E., Maria M.A., Martienssen R., McCombie W.R.;
 "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 thaliana."
 RL Nature 402:769-777 (1999).
 CC -!- FUNCTION: Might be involved in the biosynthesis of steroids
 important for the polar elongation of cells during development.
 CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum
 (Potential).
 CC -!- TISSUE SPECIFICITY: Ubiquitous.
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC -!- CAUTION: Ref.2 (CAA18139) sequence differs from that shown due to
 erroneous gene model prediction.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; ABC08097; BAA37167.1;
 CC EMBL; AL161589; CAB80304.1; ALT INIT.
 CC EMBL; Z99708; CAB16850.1; ALT INIT.
 CC EMBL; AL022141; CAA18139.1; ALT SEQ.
 CC InterPro; IPR001128; Cytochrome_P450.
 CC Pfam; PF00067; P450; 1.
 CC PRINTS; PR00385; P450.
 CC PROSITE; PS00086; CYTOCHROME_P450; 1.
 CC Oxidoreductase; Monooxygenase; Transmembrane; Heme;
 KW Endoplasmic reticulum; Multigene family.
 FT TRANSMEM 4 24 POTENTIAL.
 FT METAL 463 45
 FT CONFLICT 45 45 F -> L (IN REF. 2).
 FT SEQUENCE 524 AA; 59389 MW; 550578908BDDF272 CRC64;
 SQ
 Query Match 75.0%; Score 30; DB 1; Length 524;
 Best Local Similarity 77.8%; Pred. No. 62;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 RLSSMVKKV 9
 DB 266 RLKMKVKV 274

RESULT 10
 SNTD LUTLO STANDARD; PRT; 572 AA.

ID SNTD LUTLO STANDARD; PRT; 572 AA.

AC Q9XZ43;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 5'-nucleotidase precursor (EC 3.1.3.5).
 GN SNUC.
 OS Lutzomyia longipalpis (Sand fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Psychodidae;
 OC Psychodidae; Lutzomyia.
 OX NCBI_TaxID=7200;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Jaccobina; TISSUE=Salivary gland;
 RC Charlab R., Valenzuela J.G., Rowton E.D., Ribeiro J.M.C.;
 RT "Toward an understanding of the biochemical and pharmacological
 complexity of the saliva of a hematophagous sand fly, Lutzomyia
 longipalpis."
 RL longipalpis;
 RN [2]
 RP Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP CHARACTERIZATION.
 RC TISSUE=Salivary gland;
 RX MEDLINE=20193850; PubMed=10727894;
 RA Ribeiro J.M.C., Rowton E.D., Charlab R.;
 RA "The salivary 5'-nucleotidase/phosphodiesterase of the hematophagous

RT sand fly corrected, Lutzomyia longipalpis.";
RL Insect Biochem. Mol. Biol. 30:279-285(2000).
RN [3]
RP ERATUM.
RA Ribeiro J.M.C., Rowton E.D., Charlab R.;
RL Insect Biochem. Mol. Biol. 30:609-609(2000).
CC -|- FUNCTION: Degradation of external UDP-glucose to uridine
CC monophosphate and glucose-1-phosphate, which can then be used by
CC the cell (By similarity).
CC -|- CATALYTIC ACTIVITY: UDP-sugar + H(2)O = UMP + sugar 1-phosphate.
CC -|- CATALYTIC ACTIVITY: A 5'-ribonucleotide + H(2)O = a ribonucleoside
CC + phosphate.
CC -|- COFACTOR: Zinc. (By similarity).
CC -|- SIMILARITY: Belongs to the 5'-nucleotidase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF132510; AAD32190.1; -.
CC HSP; P07024; 2USH.
DR InterPro; IPR008334; 5'-Nucleotidase_C.
DR InterPro; IPR008146; 5'-Nucleotidase_N.
DR InterPro; IPR008179; 5'-nucleotidase.
DR InterPro; IPR004843; M-peptidase.
DR Pfam; PF02872; 5-nucleotidase; 1.
DR Pfam; PF00149; Metallophos; 1.
DR PRINTS; PR01607; APYRASEFAMLY.
DR PROSITE; PS00785; 5_NUCLEOTIDASE_1; 1.
DR PROSITE; PS00786; 5_NUCLEOTIDASE_2; 1.
KW Hydrolase; Glycoprotein; Signal; Zinc.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 572 5'-NUCLEOTIDASE.
FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 454 454 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 490 490 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 572 AA; 63353 MW; 69A652338C04536D CRC64;

Query Match 75.0%; Score 30; DB 1; Length 572;
Best Local Similarity 75.0%; Pred. NO. 67;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMWKX 8
DB 70 RVSTWVKX 77

RESULT 11
MBP1 YEAST
ID MBP1 YEAST STANDARD; PRT; 833 AA.
AC P39478.
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Transcription factor MBP1 (MBF subunit P120).
GN MBP1 OR YDL056W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OC NCBI_TaxID=4932;
OX [1]
RN SEQUENCE FROM N.A., AND SEQUENCE OF 372-387.
RP STRAIN=K1107;
RX MEDLINE=93383264; PubMed=8372350;
RA Koch C., Moll T., Neuberg M., Ahorn H., Nasmyth K.;
RT "A role for the transcription factors Mbpl and Swi4 in progression
RT from G1 to S phase.";
RL Science 261:1551-1557(1993).
RN [2]

RP SEQUENCE FROM N.A.
RA Bloecker H., Brandt P.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 460-833 FROM N.A.
RA Benton B.K., Plump S.D., Roos J., Lennarz W.J., Cross F.R.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-100.
RX MEDLINE=97446148; PubMed=9299332;
RA Taylor I.A., Treiber M.K., Olivi L., Smerdon S.J.;
RT "The X-ray structure of the DNA-binding domain from the Saccharomyces
RT cerevisiae cell-cycle transcription factor Mbpl at 2.1-A
RT resolution.";
RL J. Mol. Biol. 272:1-8(1997).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.71 ANGSTROMS) OF 1-102.
RX MEDLINE=97238931; PubMed=9083114;
RA Xu R.M., Koch C., Liu Y., Horton J.R., Knapp D., Nasmyth K., Cheng X.;
RT "Crystal structure of the DNA-binding domain of Mbpl, a transcription
RT factor important in cell-cycle control of DNA synthesis.";
RL Structure 5:349-358(1997).
CC -|- FUNCTION: Binds to MCB elements (Mlu I cell cycle box) found in
CC the promoter of most DNA synthesis genes. Transcriptional
CC activation by MBF has an important role in the transition from G1
CC to S phase. It may have a dual role in that it behaves as an
CC activator of transcription at the G1-S boundary and as a repressor
CC during other stages of the cell cycle.
CC -|- SUBUNIT: MBF contains Swi6 and MBP1.
CC -|- SUBCELLULAR LOCATION: Nuclear.
CC -|- SIMILARITY: Contains 2 ANK repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X74158; CAAS2271.1; -.
CC EMBL; Z74104; CAA98618.1; -.
CC EMBL; U19608; AAC49290.1; -.
CC PIR; A47528; A47528.
CC PDB; 1MBL; 29-JUL-98.
CC PDB; 1MB8; 02-MAR-99.
CC PDB; 1L3G; 18-FEB-03.
CC GeneOnline; 140298; -.
CC TRANSFAC; T03480; -.
CC SGD; S0002214; MBP1.
CC GO; GO:0005634; C:nucleus; IDA.
CC GO; GO:0003700; F:transcription factor activity; IMP.
CC GO; GO:0006260; P:DNA replication; IMP.
CC GO; GO:0000074; P:regulation of cell cycle; IGI.
CC InterPro; IPR02110; ANK.
CC InterPro; IPR003163; Yeast_DNA_bind.
CC Pfam; PF00023; ank; 3.
CC Pfam; PF02292; APSES; 1.
CC SMART; SM0248; ANK; 2.
CC PROSITE; PS50089; ANK_REPEAT; 2.
CC PROSITE; PS50297; ANK_REPEAT_REGION; 1.
CC Transcription regulation; Activator; Nuclear protein; DNA-binding;
CC ANK repeat; Repeat; 94
FT DNA BIND 1 94
FT REPEAT 394 423 ANK 1.
FT REPEAT 512 541 ANK 2.
FT STRAND 5 10
FT TURN 11 12
FT STRAND 13 19
FT STRAND 24 28
FT TURN 29 31
FT STRAND 34 35
FT STRAND 36 42
FT HELIX

```
FT TURN 43 44
FT HELIX 47 57
FT TURN 58 60
FT STRAND 64 66
FT TURN 72 73
FT STRAND 75 78
FT HELIX 80 89
FT TURN 90 91
FT HELIX 93 101
SQ SEQUENCE 833 AA; 93907 MW; BB7C35E29802EBD5 CRC64;

Query Match 75.0%; Score 30; DB 1; Length 833;
Best Local Similarity 75.0%; Pred. No. 95;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKK 8
DB 767 KLSSLVKK 774

RESULT 12
T2D7 RAT STANDARD; PRT; 253 AA.
AC Q62880;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Transcription initiation factor TPIID 31 kDa subunit (TAFII-31)
DE (TAFII-32) (Neuronal cell death related gene in neuron -7)
DE (DN-7)
GN TAF9 OR TAF2G OR TAFII31.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=971253; PubMed=9168994;
RA Aoki T., Koike T., Nakano T., Shibahara K., Nishimura H., Kikuchi H.,
RA Honjo T.;
RT "Rat TAFII31 gene is induced upon programmed cell death in
RT differentiated PC12 cells deprived of NGF.";
RL Biochem. Biophys. Res. Commun. 234:230-234(1997).
CC -!- FUNCTION: TAFs are components of the transcription factor IID
CC (TFIID) complex that are essential for mediating regulation of RNA
CC polymerase transcription. TAFII31 is a coactivator for the p53
CC protein. Also interacts with the acidic transcription factor
CC protein 16 (VP16) as well as with the general transcription factor
CC TFIIB (By similarity).
CC -!- SUBUNIT: TPIID is composed of TATA binding protein (TBP) and a
CC number of TBP-associated factors (TAFs). TAFII31 and TAF33 bind to
CC each other via amino acid residues in the amino-terminal domain of
CC TAF33 that are essential for transcription. Component of the PCAF
CC complex, at least composed of TADA2L/ADA2, TADA3L/ADA3, SUPT3H,
CC TAF5L/PAF5beta, TAF6L/PAF6alpha, TAF10/TAFII30, TAF12/TAFII20,
CC TAF9/TAFII31 and TRRAP. Component of the STAGA transcription
CC coactivator-HAT complex, at least composed of SF3B3/SAP130,
CC GCN5L2, STAF65gamma/XIRA0764, TAF5L, TAF6L, TADA3L, TAF10, TAF12,
CC TRAP and TAF9 (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the TAF3G family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; U40186; AAC53201.1; -
PIR; JC5511; JC5511.
InterPro; IPR007124; Hist_TAF.

DR InterPro; IPR003162; TFIID-31.
DR Pfam; PF02291; TFIID-31; 1.
DR ProDom; PD011023; TFIID-31; 1.
KW Transcription regulation; Nuclear protein.
FT DOMAIN 238 251 POLY-ASP.
SQ SEQUENCE 253 AA; 27620 MW; 42045091A9B94378 CRC64;

Query Match 72.5%; Score 29; DB 1; Length 253;
Best Local Similarity 75.0%; Pred. No. 53;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMVKK 8
DB 123 RLKSLVKK 130

RESULT 13
EGC2 OCEIH STANDARD; PRT; 351 AA.
AC Q8ETB7;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Probable GTPase engC protein 2 (EC 3.6.1.-).
DE ENG2 OR OBO344.
OS Oceanobacillus iheyensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
OX NCBI_TaxID=182710;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=HTE831 / DSM 14371 / JCM 11309;
RX MEDLINE=2220767; PubMed=12235376;
RA Takami H., Takaki Y., Uchiyama I.;
RA "Genome sequence of Oceanobacillus iheyensis isolated from the Iheya
RA Ridge and its unexpected adaptive capabilities to extreme
RA environments.";
RL Nucleic Acids Res. 30:3927-3935(2002).
CC -!- FUNCTION: Unusual circularly permuted GTPase that catalyzes rapid
CC hydrolysis of GTP with a slow catalytic turnover (By similarity).
CC -!- SUBUNIT: Monomer (Probable).
CC -!- SIMILARITY: Contains 1 engC GTPase domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; AP004594; BAC12300.1; -
HAMAP; MF_01820; -; 1.
DR InterPro; IPR004881; DUF258.
DR Pfam; PF03193; DUF258; 1.
DR TIGRFAMs; TIGR00157; TIGR00157; 1.
DR PROSITE; PS00936; ENG_C_GTPASE; 1.
KW Hydrolase; GTP-binding; Complete protease.
FT DOMAIN 109 256 ENG_C_GTPASE.
FT NP_BIND 148 151 GTP (PROBABLE).
FT NP_BIND 200 207 GTP (PROBABLE).
FT NP_BIND 251 255 GTP (PROBABLE).
FT SITE 282 295 KNUCKLE-LIKE CYSTEINE CLUSTER.
SQ SEQUENCE 351 AA; 40008 MW; E0DA5E8EDCD57FF7 CRC64;

Query Match 72.5%; Score 29; DB 1; Length 351;
Best Local Similarity 66.7%; Pred. No. 72;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKK 9
DB 89 RFSIVRKV 97
```

RESULT 14

CRTI_MAIZE
 ID CRTI_MAIZE STANDARD; PRT; 571 AA.
 AC P49086; Q41849;
 DT 01-REB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Phytoene dehydrogenase, chloroplast precursor (EC 1.14.99.-)
 DE (Phytoene desaturase).
 GN PDS1 OR PDS.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACMAD clade; Panicoideae; Andropogoneae; Zea.
 NCBI_taxid=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96178866; PubMed=8616251;
 RA Li Z., Matthews P.D., Burr B., Wurtzel E.T.,
 RT "Cloning and characterization of a maize cDNA encoding phytoene
 RT desaturase, an enzyme of the carotenoid biosynthetic pathway."
 RL Plant Mol. Biol. 30:269-279(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=cv. Funkf; TISSUE=Leaf;
 RA MEDLINE=95357435; PubMed=7630964;
 RA Habbe W.E., Oishi K.K.;
 RT "Maize phytoene desaturase maps near the viviparous5 locus."
 RL Plant Physiol. 108:1329-1330(1995).
 CC -!- FUNCTION: This enzyme converts phytoene into zeta-carotene via the
 CC intermediary of phytofluene by the symmetrical introduction of two
 CC double bonds at the C-11 and C-11' positions of phytoene.
 CC -!- COFACTOR: NAD, NADP, or FAD (Probable).
 CC -!- PATHWAY: Carotenoid biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Chloroplast; chromoplast.
 CC -!- SIMILARITY: Belongs to the phytoene dehydrogenase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; U37285; AAC12846.1; -
 CC FIRM; L39286; AAA95919.1; -
 CC FIRM; S65060; S65060.
 CC MaizeDB; 84977; -
 CC InterPro; IPR001613; Amineoxid fl.
 CC InterPro; IPR002937; Amino oxidase.
 CC Pfam; PF01593; Amino oxidase; 1.
 CC PRINTS; PR00757; AMINEOXDASEF.
 CC Carotenoid biosynthesis; Oxidoreductase; NAD; Flavoprotein; FAD;
 CC Chloroplast; transit peptide.
 CC TRANSIT 1 96 CHLOROPLAST (POTENTIAL).
 CC CHAIN 97 571 PHYTOENE DEHYDROGENASE.
 CC CONFLICT 61 63 VVC -> LSA (IN REF. 2).
 CC CONFLICT 68 68 R -> S (IN REF. 2).
 CC CONFLICT 555 555 A -> T (IN REF. 2).
 CC SEQUENCE 571 AA; 64115 MW; FAEL19C78FBE799A CRC64;
 Query Match 72.5%; Score 29; DB 1; Length 571;
 Best Local Similarity 55.6%; Pred. No. 1.1e+02;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 RLSSWVKV 9
 Db 329 RLNSRIKKI 337

RESULT 15

CRTI_CAPAN
 ID CRTI_CAPAN STANDARD; PRT; 582 AA.
 AC P80093;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Phytoene dehydrogenase, chloroplast precursor (EC 1.14.99.-)
 DE (Phytoene desaturase).
 GN PDS.
 OS Capsicum annuum (Bell pepper).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; asteride;
 OC Lamiales; Solanales; Solanaceae; Capsicum.
 NCBI_taxid=4072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=cv. Lamuyo;
 RX MEDLINE=93011154; PubMed=1396714;
 RA Huguency P., Roemer S., Kuntz M., Camara B.;
 RT "Characterization and molecular cloning of a flavoprotein catalyzing
 RT the synthesis of phytofluene and zeta-carotene in Capsicum
 RT chromoplasts."
 RL Eur. J. Biochem. 209:399-407(1992).
 CC -!- FUNCTION: This enzyme converts phytoene into zeta-carotene via the
 CC intermediary of phytofluene by the symmetrical introduction of two
 CC double bonds at the C-11 and C-11' positions of phytoene.
 CC -!- COFACTOR: NAD, NADP, or FAD (Probable).
 CC -!- ENZYME REGULATION: Inhibited by the herbicides metflurazon,
 CC diflufenone, fluridone and diflufenican.
 CC -!- PATHWAY: Carotenoid biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Chloroplast; chromoplast.
 CC -!- DEVELOPMENTAL STAGE: Ripening fruit.
 CC -!- SIMILARITY: Belongs to the phytoene dehydrogenase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; X68058; CAA48195.1; -
 CC FIRM; S29314; S29314.
 CC InterPro; IPR001613; Amineoxid fl.
 CC InterPro; IPR002937; Amino oxidase.
 CC Pfam; PF01593; Amino oxidase; 1.
 CC PRINTS; PR00757; AMINEOXDASEF.
 CC Carotenoid biosynthesis; Oxidoreductase; NAD; Flavoprotein; FAD;
 CC Chloroplast; transit peptide; Herbicide resistance.
 CC TRANSIT 1 110 CHLOROPLAST (POTENTIAL).
 CC CHAIN 111 582 PHYTOENE DEHYDROGENASE.
 CC NP_BIND 117 133 FAD (ADP PART) (POTENTIAL).
 CC SEQUENCE 582 AA; 65061 MW; 011E6E7DCAF3DB5 CRC64;
 Query Match 72.5%; Score 29; DB 1; Length 582;
 Best Local Similarity 55.6%; Pred. No. 1.1e+02;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 RLSSWVKV 9
 Db 343 RLNSRIKKI 351

Search completed: March 1, 2004, 17:29:55
 Job time : 8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:20:41 ; Search time 30.3333 Seconds

(without alignments)
93.615 Million cell updates/sec

Title: US-09-905-083-32

Perfect score: 40

Sequence: 1 RLSSMVKKV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL 25:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp Vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	181	4 Q8NFV7	Q8nfv7 homo sapien
2	40	100.0	253	4 Q8NSN9	Q8nsn9 homo sapien
3	33	82.5	282	16 Q8S2B8	Q8s2b8 streptomyce
4	33	82.5	417	5 Q8WTE7	Q8wte7 anopheles g
5	32	80.0	169	16 Q899C6	Q899c6 clostridium
6	32	80.0	351	17 Q97Z26	Q97z26 sulfolobus
7	32	80.0	664	3 Q9F7W7	Q9f7w7 schizosacch
8	31	77.5	97	9 Q8B021	Q8b021 bacterioph
9	31	77.5	390	10 Q8LEQ6	Q8leq6 arabidopsis
10	31	77.5	550	9 Q8W607	Q8w607 bacterioph
11	31	77.5	589	10 Q7XX91	Q7xx91 oryza sativ
12	31	77.5	610	16 Q8EYH8	Q8eyh8 leptospira
13	31	77.5	751	10 Q7XNM7	Q7xnm7 oryza sativ
14	31	77.5	912	10 Q9LKD6	Q9ldk6 arabidopsis
15	31	77.5	2159	10 Q8RVL1	Q8rvl1 zea mays (m
16	31	77.5	2159	10 Q8RUQ1	Q8ruq1 zea mays (m

17	30	75.0	105	10 Q93W83	Q93w83 arabidopsis
18	30	75.0	124	11 Q8BP39	Q8bp39 mus musculus
19	30	75.0	131	10 Q851X1	Q851x1 oryza sativ
20	30	75.0	236	10 Q80845	Q80845 arabidopsis
21	30	75.0	240	5 Q17203	Q17203 caenorhabdi
22	30	75.0	308	5 Q86HA4	Q86ha4 dictyosteli
23	30	75.0	320	11 Q8BP92	Q8bp92 mus musculus
24	30	75.0	321	11 Q70341	Q70341 mus musculus
25	30	75.0	348	2 Q8KNA1	Q8kna1 pseudomonas
26	30	75.0	373	17 Q9V014	Q9v014 pyrococcus
27	30	75.0	391	10 Q9C6U9	Q9c6u9 arabidopsis
28	30	75.0	394	10 Q9C857	Q9c857 arabidopsis
29	30	75.0	517	10 Q94C74	Q94c74 arabidopsis
30	30	75.0	533	10 Q8GK11	Q8gk11 arabidopsis
31	30	75.0	578	5 Q86PF2	Q86pf2 drosophila
32	30	75.0	647	2 Q49468	Q49468 mycoplasma
33	30	75.0	661	13 Q7T3S2	Q7t3s2 xenopus lae
34	30	75.0	690	16 Q7UQT1	Q7uqt1 rhodospirell
35	30	75.0	815	3 Q9HEF3	Q9hef3 neurospora
36	30	75.0	986	17 Q9UXE1	Q9uxb1 sulfolobus
37	30	75.0	1167	16 Q9PAG0	Q9pag0 xylella fas
38	30	75.0	1167	16 Q87A86	Q87a86 xylella fas
39	30	75.0	1768	5 Q9N8K7	Q9n8k7 trypanosoma
40	30	75.0	1887	5 Q9GZG6	Q9gzg6 caenorhabdi
41	29	72.5	99	16 Q7VA27	Q7va27 prochloroco
42	29	72.5	118	10 Q40388	Q40388 nicotiana b
43	29	72.5	119	6 Q8HXH8	Q8hxs8 macaca faec
44	29	72.5	136	10 Q7XZC6	Q7xz26 nicotiana b
45	29	72.5	208	10 Q8L418	Q8l418 oryza sativ

ALIGNMENTS

RESULT 1

Q8NFV7	PRELIMINARY;	PRT;	181 AA.
ID	Q8NFV7		
AC	Q8NFV7;		
DT	01-OCT-2002 (TrEMBLrel. 22, Created)		
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	Kallikrein 7 short variant protein.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]_TaxID=9606;		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Ovarian carcinoma;		
RA	Dong Y., Kaushal A., Clements J.A.;		
RT	"Human Kallikrein 7 (KLK7) short variant mRNA from ovarian carcinoma."		
RL	Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.		
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.		
DR	EMBL, AF411215; AN03663.1; -		
DR	GO; GO:0004263; F:chymotrypsin activity; IEA.		
DR	GO; GO:0008233; F:peptidase activity; IEA.		
DR	GO; GO:0004295; F:trypsin activity; IEA.		
DR	GO; GO:0008508; P:proteolysis and peptidolysis; IEA.		
DR	InterPro; IPR009003; Cys Ser trypsin.		
DR	InterPro; IPR001254; Peptidase S1.		
DR	PFam; PF00889; Trypsin; 1.		
DR	PRINTS; PR00722; CHYMOTRYPsin.		
DR	SMART; SM00020; TRYP SPc; 1.		
DR	PROSITE; PS0240; TRYPsin DOM; 1.		
DR	PROSITE; PS00135; TRYPsin SER; 1.		
KW	Hydrolase; Protease; Serine protease.		
SQ	SEQUENCE 181 AA; 19887 MW; 86A28A03B80C2D78 CRC64;		

Query Match 100.0%; Score 40; DB 4; Length 181;

Best Local Similarity 100.0%; Pred. No. 0.86; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0;

```

QY 1 RLSSWVKV 9
DB 51 RLSSWVKV 59

RESULT 2
ID Q8NSN9 PRELIMINARY; PRT; 253 AA.
AC Q8NSN9;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Kallikrein 7 (chymotryptic, stratum corneum).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; BC032005; AAH32005.1; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0006233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; tryp_SPC; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 253 AA; 27608 MW; 2D68B6A41B22A668 CRC64;

Query Match 100.0%; Score 40; DB 4; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.2; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

QY 1 RLSSWVKV 9
DB 123 RLSSWVKV 131

RESULT 3
ID Q9S2Q8 PRELIMINARY; PRT; 282 AA.
AC Q9S2Q8;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Putative endo alpha-1,4 polygalactosaminidase.
GN SC02178 OR SC5F7.23C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidaigo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,

```

```

RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939111; CAB51262.1; -.
DR PIR; T3294; T35294.
DR InterPro; IPR000437; Prok_lipoprot_S.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Complete proteome.
SQ SEQUENCE 282 AA; 30789 MW; BEECFE1743703A33 CRC64;

Query Match 82.5%; Score 33; DB 16; Length 282;
Best Local Similarity 87.5%; Pred. No. 47;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSWVKV 8
DB 265 RLSSWVKV 272

RESULT 4
ID Q8WTE7 PRELIMINARY; PRT; 417 AA.
AC Q8WTE7;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Putative odorant receptor Ori.
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21593012; PubMed=11724964;
RA Fox A.N., Pitts R.J., Robertson H., Carlson J., Zwiebel L.J.;
RT "Anopheles gambiae Putative Odorant Receptors.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:14693-14697(2001).
DR EMBL; AF364130; AAL35506.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005549; F:odorant binding; IEA.
DR GO; GO:0004984; F:olfactory receptor activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007608; P:olfaction; IEA.
DR InterPro; IPR004117; 7tm_6.
DR Pfam; PF02949; 7tm_6; 1.
KW Receptor.
SQ SEQUENCE 417 AA; 48520 MW; F60D7D7BD93D37F2 CRC64;

Query Match 82.5%; Score 33; DB 5; Length 417;
Best Local Similarity 77.8%; Pred. No. 68;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSWVKV 9
DB 219 RLSSWVKV 227

RESULT 5
ID Q899C6 PRELIMINARY; PRT; 169 AA.
AC Q899C6;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Anaerobic ribonucleoside-triphosphate reductase activating protein (EC
DE 1.97.1.1-).
GN CTC00258.
OS Clostridium tetani.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.

```

```

OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Massachusetts / E88;
RX MEDLINE=22457253; PubMed=12552129;
RA Brueggemann H., Baumer S., Fricke W.F., Wieser A., Liesegang H.,
RA Decker I., Hertzberg C., Martinez-Arias R., Merkl R., Henne A.,
RA Gottschalk G.;
RT "The genome sequence of Clostridium tetani, the causative agent of
tetanus disease.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).
DR EMBL; AE015936; AAC34903.1; -
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; F:electron transport; IEA.
DR GO; GO:0006800; P:oxygen and reactive oxygen species metabolism; IEA.
DR InterPro; IPR000345; CytC heme BS.
DR InterPro; IPR001989; Radical activat.
DR InterPro; IPR007197; Radical SAM.
DR Pfam; PF04055; Radical SAM; I.
DR PROSITE; PS00190; CYTOCHROME C; 1.
DR PROSITE; PS01087; RADICAL_ACTIVATING; 1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 169 AA; 19292 MW; CC7AB2511D3844C1 CRC64;

Query Match 80.0%; Score 32; DB 16; Length 169;
Best Local Similarity 75.0%; Pred. No. 49;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVVK 8
DB 85 RLASMIXK 92
|||:|||||
|||:|||||

RESULT 6
Q97226 PRELIMINARY; PRT; 351 AA.
AC Q97226;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Biotin synthase (biob) (EC 2.8.1.6).
GN BI0B OR S80115
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=23332296; PubMed=11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aways M.J., Chan-Waiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Erasuo G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Regan M.A., Senses C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL; AB006729; AAK41371.1; -
DR PIR; D90264; D90264.
DR GO; GO:0004076; F:biotin synthase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR006638; Elp3.
DR InterPro; IPR007197; Radical SAM.
DR Pfam; PF04055; Radical SAM; 1.
DR SMART; SM00729; Elp3; 1.
KW Transferase; Complete proteome.
SQ SEQUENCE 351 AA; 39814 MW; 892DB90BC987F752 CRC64;

Query Match 80.0%; Score 32; DB 17; Length 351;
Best Local Similarity 75.0%; Pred. No. 49;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

us-09-905-083-32.rspt

Best Local Similarity 66.7%; Pred. No. 97;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMVVKV 9
DB 341 RLDSLKKV 349
|||:|||||
|||:|||||

RESULT 7
Q9P7W7 PRELIMINARY; PRT; 664 AA.
AC Q9P7W7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical Armadillo/beta-catenin domain protein.
GN SPBCL703.03C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA McDougall R.C., Rajandream M.A., Barrell B.G., Cadieu E., Lelaure V.,
RA Galibert F.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL136536; CAB66447.1; -
DR PIR; T50316; T50316.
DR GenDB SPombe; SPBCL703.03C; -
DR InterPro; IPR000225; Armadillo.
DR Pfam; PF00514; Armadillo_seg; 1.
DR PROSITE; PS50176; ARM_REPEAT; 1.
SQ SEQUENCE 664 AA; 74811 MW; 363FEE00EB15569F CRC64;

Query Match 80.0%; Score 32; DB 3; Length 664;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMVVKV 9
DB 624 RLKNVKKI 632
|||:|||||
|||:|||||

RESULT 8
Q9B021 PRELIMINARY; PRT; 97 AA.
AC Q9B021;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Bacteriophage GMS5-1.
OC Viruses.
OX NCBI_TaxID=148339;
RN [1]
RP SEQUENCE FROM N.A.
RA Dale C., Young S.A., Maudlin I., Welburn S.C.;
RT "Endosymbiont bacteriophage may influence susceptibility to
trypanosome infection in tsetse".
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF111657; AAG50264.1; -
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 97 AA; 10987 MW; B17CB63BE43F2C52 CRC64;

Query Match 77.5%; Score 31; DB 9; Length 97;
Best Local Similarity 87.5%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LSSMVKKV 9
|||:|||||
|||:|||||
```

```

Db      88 LQSMVKV 95
RESULT 9
Q8LEQ6 PRELIMINARY; PRT; 390 AA.
AC Q8LEQ6;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Haas B.J., Volkovskiy N., Town C.D., Troughan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation.";
RL Genome Biol. 0:0-0(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Proyer V., Troughan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY085293; AR062525.1; -
DR InterPro; IPR001810; F-box.
DR InterPro; IPR006652; Kelch_rep.
DR Pfam; PF00546; F-box; 1.
DR Pfam; PF01344; Kelch; 1.
DR SMART; SM00256; Fbox; 1.
DR Hypothetical protein.
KW SEQUENCE 390 AA; 44992 MW; 2ABA01DDF4E19F72 CRC64;

Query Match 77.5%; Score 31; DB 10; Length 390;
Best Local Similarity 55.6%; Pred. No. 1.8e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LQSMVKV 9
Db 230 RFSSMIRKI 238

RESULT 10
Q8W607 PRELIMINARY; PRT; 550 AA.
AC Q8W607;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Terminusase large subunit.
GN TERL.
OS Bacteriophage PSA.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=171618;
RN [1]
RP SEQUENCE FROM N.A.
RA Leesner M.J., Sattelberger E., Zimmer M., Calendar R., Inman R.B.,
RA Scherzer S.;
RT Molecular analysis of Listeria monocytogenes ScottA bacteriophage PSA
RT reveals ribosomal frameshifting as a general mechanism for generation
RT of major structural proteins.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ312240; CAC8558.1; -
DR InterPro; IPR005021; Phage termin.
DR Pfam; PF03354; Phage terminase; 1.
DR Pfam; PF03354; Phage terminase; 1.
KW SEQUENCE 550 AA; 63915 MW; 0F43142D3906F00F CRC64;

Query Match 77.5%; Score 31; DB 9; Length 550;

Best Local Similarity 77.8%; Pred. No. 2.5e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSWVKV 9
Db 469 RLSSWVDKL 477

RESULT 11
Q7XX91 PRELIMINARY; PRT; 589 AA.
AC Q7XX91;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE OSJNB0074B10.5 protein.
GN OSJNB0074B10.5.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Fu G., Wang S.-Y., Ren S.-X., Lv G., Lin W., Gu W.-Q., Zhu G.-F., Tu Y.-F.,
RA Jia J., Yin H.-F., Zhang Y., Cai Z., Chen J., Kang H., Chen X.-Y.,
RA Shao Y., Sun Y., Hu Q.-P., Zhang X.-L., Zhang W., Wang L.-J., Ding C.-W.,
RA Sheng H.-H., Gu J.-L., Chen S.-T., Ni L., Zhu F.-H., Han B., Feng Q.,
RA Huang Y.-C., Li Y., Zhu J.-J., Zhao Q., Hu X., Liu Y.-L., Mu J., Yu Z.,
RA Chen L., Fan D.-L., Weng Q.-J., Zhang L., Lu Y.-Q., Yu S.-L., Liu X.-H.,
RA Lu T.-T., Zhang Y.-J., Lu Y., Li C., Li T., Zhang Y., Hu H., Jia P.-X.,
RA Qian Y.-M., Ying K., Zhou B., Chen Z.-H., Hao P., Zhang L., Wu M.,
RA Zhang R.-Q., Guan J.-P., Hong G.-F.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ662941; CAD39649.1; -
SQ SEQUENCE 589 AA; 68068 MW; 2673B44619A9EB74 CRC64;

Query Match 77.5%; Score 31; DB 10; Length 589;
Best Local Similarity 87.5%; Pred. No. 2.6e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSSWVKV 9
Db 374 LSSWVKV 381

RESULT 12
Q8EYH8 PRELIMINARY; PRT; 610 AA.
AC Q8EYH8;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Mts-like mismatch repair protein, ATPases.
GN MUTS3 OR LA4236.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011576; AAN51434.1; -
DR GO; GO:0005524; P-ATP binding; IEA.
DR GO; GO:0003684; F-damaged DNA binding; IEA.
DR GO; GO:0006298; P-mismatch repair; IEA.
DR InterPro; IPR000432; MUTS_C.
DR Pfam; PF00488; Muts V; 1.
DR ProDom; PD001263; Muts C; 1.
DR SMART; SM00534; MUTSac; 1.
DR Complete proteome.
KW SEQUENCE 610 AA; 70754 MW; 08E16897C1D731E5 CRC64;

```


Query Match 77.5%; Score 31; DB 16; Length 610;
 Best Local Similarity 56.7%; Pred. No. 2.7e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9
 |||:|:|:
 DB 502 RLSEIVKKI 510

RESULT 13
 Q7XNM7 PRELIMINARY; PRT; 751 AA.
 AC Q7XNM7; 751 AA.
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE OSJNB0076A11.5 protein.
 GN OSJNB0076A11.5
 OS Oryza sativa (Rice)
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OC NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
 RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,
 RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,
 RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,
 RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
 RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
 RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
 RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
 RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.
 RL Submitted (SFP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL060693; CA04521.1; -
 SQ SEQUENCE 751 AA; 87386 MW; 7F47A64A44A1V182C CRC64;

Query Match 77.5%; Score 31; DB 10; Length 751;
 Best Local Similarity 87.5%; Pred. No. 3.3e+02;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSSMVKKV 9
 |||:|:|:
 DB 404 LSSVVKKV 411

RESULT 14
 Q9LDK6 PRELIMINARY; PRT; 912 AA.
 AC Q9LDK6; 912 AA.
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN A74339750.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Monfort A., Casacuberta E., Puigdomenech P., Mewes H.W., Lemcke K.,
 RA Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.

RA Bevan M., Monfort A., Casacuberta E., Puigdomenech P., Hoheisel J.,
 RA Mewes H.W., Lemcke K., Mayer K.F.X.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL161595; CAB80638.1; -
 DR EMBL; AL022605; CAB77062.1; -
 DR PIR; H85470; H85470.
 DR PIR; T05012; T05012.
 DR InterPro; IPR001810; F-box.
 DR InterPro; IPR006852; Kelch_rep.
 DR Pfam; PF00646; F-box; 3.
 DR Pfam; PF01344; Kelch; 2.
 DR SMART; SM00256; FBOX; 2.
 DR SMART; SM00612; Kelch; 2.
 DR PROSITE; PS00181; FBOX; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 912 AA; 105278 MW; 8C764AFDD114D57E CRC64;

Query Match 77.5%; Score 31; DB 10; Length 912;
 Best Local Similarity 55.6%; Pred. No. 4e+02;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9
 |||:|:|:
 DB 403 RFSMIRKI 411

RESULT 15
 Q8RVLL PRELIMINARY; PRT; 2159 AA.
 ID Q8RVLL
 AC Q8RVLL
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Calpain-like protein.
 GN DEK1.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACAD clade; Panicoideae; Andropogoneae; Zea.
 OC NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.

RA STRAIN=cv. B73; TISSUE=Developing endosperm;
 RC PubMed=11929961;
 RX Lid S.E., Grulis D., Jung R., Lorentzen J.A., Ananiev E.,
 RA Chamberlin M., Niu X., Mealey R., Nichols S., Olsen O.A.; cell
 RT "The defective kernel 1 (dek1) gene required for aleurone cell
 RT development in the endosperm of maize grains encodes a membrane
 RT protein of the calpain gene superfamily."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:5460-5465(2002).
 DR EMBL; AY061806; NAL38189.1; -
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0004198; F:calpain activity; IEA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
 DR GO; GO:0006306; P:DNA methylation; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR002052; N6.Mtase.
 DR InterPro; IPR001300; Peptidase_C2.
 DR InterPro; IPR000189; SHprot_acsite.
 DR Pfam; PF01067; Calpain_III; 1.
 DR Pfam; PF00648; Peptidase_C2; 1.
 DR SMART; SM00720; calpain_III; 1.
 DR SMART; SM00230; CysPC; 1.
 DR PROSITE; PS00092; N6.MTASE; 1.
 DR PROSITE; PS00139; THIOI.PROTEASE.CYS; 1.
 SQ SEQUENCE 2159 AA; 238997 MW; 9F995D3F9EAC1315 CRC64;

Query Match 77.5%; Score 31; DB 10; Length 2159;
 Best Local Similarity 75.0%; Pred. No. 8.9e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 8

Db 515 RISSMLXX 522
|:|||||

Search completed: March 1, 2004, 17:34:42
Job time : 33.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:21:46 ; Search time 11.8899 Seconds
(without alignments)
39.081 Million cell updates/sec

Title: US-09-905-083-33

Perfect score: 40

Sequence: 1 LLLPLQLLL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	9	3	US-09-502-600-33
2	40	100.0	9	4	US-09-918-243-33
3	40	100.0	253	2	US-08-557-146-2
4	40	100.0	253	2	US-08-824-874-3
5	40	100.0	253	2	US-09-154-344-2
6	40	100.0	253	3	US-08-930-188-2
7	40	100.0	253	3	US-09-210-084-3
8	40	100.0	253	4	US-09-764-762-3
9	40	100.0	253	5	PCT-US96-04294-2
10	38	95.0	812	4	US-09-489-039A-12075
11	36	90.0	9	3	US-09-502-600-35
12	36	90.0	9	3	US-09-502-600-36
13	36	90.0	9	4	US-09-918-243-35
14	36	90.0	9	4	US-09-918-243-36
15	33	82.5	190	1	US-08-339-152A-19
16	33	82.5	190	2	US-08-007-999B-6
17	33	82.5	190	2	US-08-689-276A-6
18	33	82.5	447	4	US-09-252-991A-24312
19	33	82.5	634	1	US-08-339-152A-17
20	33	82.5	653	1	US-08-339-152A-16
21	33	82.5	653	2	US-08-007-999B-3
22	33	82.5	653	2	US-08-689-276A-3
23	32	80.0	21	3	US-08-753-007A-24
24	32	80.0	21	3	US-09-398-496-24
25	32	80.0	22	4	US-08-977-378-22
26	32	80.0	235	1	US-07-940-605A-12
27	32	80.0	235	2	US-08-690-096-12

ALIGNMENTS

RESULT 1

US-09-502-600-33

; Sequence 33, Application US/09502600A

; Patent No. 6294344

; GENERAL INFORMATION:

; APPLICANT: O'Brien, Timothy J.

; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of

; FILE REFERENCE: D6223CIP-C

; CURRENT FILING DATE: 2000-02-11

; CURRENT APPLICATION NUMBER: US/09/502,600A

; PRIOR APPLICATION NUMBER: 09/039,211

; PRIOR FILING DATE: 03-14-1998

; NUMBER OF SEQ ID NOS: 136

; SEQ ID NO 33

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: Residues 5-13 of the SCCE protein

US-09-502-600-33

Query Match 100.0%; Score 40; DB 3; Length 9;

Best Local Similarity 100.0%; Pred. No. 3e+05;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQLLL 9

Db 1 LLLPLQLLL 9

RESULT 2

US-09-918-243-33

; Sequence 33, Application US/09918243

; Patent No. 6627403

; GENERAL INFORMATION:

; APPLICANT: O'Brien, Timothy J.

; APPLICANT: Cannon, Martin J.

; APPLICANT: Santin, Alessandro

; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer

; FILE REFERENCE: D6223CIP/C/D/CIP

; CURRENT APPLICATION NUMBER: US/09/918,243

; CURRENT FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: US

; PRIOR FILING DATE: 2001-07-13

; NUMBER OF SEQ ID NOS: 136

; SEQ ID NO 33

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Homo sapiens

Sequence 37, Appli
Sequence 8, Appli
Sequence 11785, A
Sequence 4, Appli
Sequence 4, Appli
Sequence 12646, A
Sequence 7, Appli
Sequence 3902, Ap
Sequence 1, Appli
Sequence 1, Appli
Sequence 8568, Ap
Sequence 5754, Ap
Sequence 3, Appli
Sequence 6, Appli
Sequence 2, Appli
Sequence 285, App
Sequence 285, App

28 32 80.0 302 4 US-09-393-634-37
29 32 80.0 376 3 US-08-751-512-8
30 32 80.0 436 4 US-09-489-039A-11785
31 32 80.0 492 1 US-07-794-393-4
32 32 80.0 492 1 US-08-001-711-4
33 32 80.0 837 4 US-09-489-039A-12646
34 32 80.0 1306 3 US-08-983-299-7
35 32 80.0 1306 4 US-09-407-427-7
36 31 77.5 134 4 US-09-621-976-3902
37 31 77.5 270 2 US-08-773-368-1
38 31 77.5 270 3 US-09-199-887-1
39 31 77.5 324 4 US-09-489-039A-8568
40 31 77.5 369 4 US-09-107-532A-5754
41 31 77.5 433 4 US-09-705-448-3
42 31 77.5 445 3 US-08-974-691-6
43 31 77.5 451 3 US-08-974-691-2
44 31 77.5 463 4 US-09-907-794A-285
45 31 77.5 463 4 US-09-905-125A-285

FEATURE:
NAME/KEY: CHAIN
OTHER INFORMATION: Residues 5-13 of the SCCE protein
US-09-918-243-33

Query Match 100.0%; Score 40; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e-05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9
|||||
Db 1 LLLPLQILL 9

RESULT 3
US-08-557-146-2
Sequence 2, Application US/08557146
Patent No. 5834290
GENERAL INFORMATION:
APPLICANT: Egelrud, Torbjorn
APPLICANT: Hansson, Lennart
TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
TITLE OF INVENTION: Enzyme (SCCE)
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: White & Case, Patent Department
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/557,146
FILING DATE: 14-DEC-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sterner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-181
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-557-146-2

Query Match 100.0%; Score 40; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9
|||||
Db 5 LLLPLQILL 13

RESULT 4
US-08-824-874-3
Sequence 3, Application US/08824874
Patent No. 5962300
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
TITLE OF INVENTION: NOVEL KALLIKREIN
NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/824,874
FILING DATE: Filed Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0252 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4186
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 532504
US-08-824-874-3

Query Match 100.0%; Score 40; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9
|||||
Db 5 LLLPLQILL 13

RESULT 5
US-09-154-344-2
Sequence 2, Application US/09154344
Patent No. 5981256
GENERAL INFORMATION:
APPLICANT: Egelrud, Torbjorn
APPLICANT: Hansson, Lennart
TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
TITLE OF INVENTION: Enzyme (SCCE)
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: White & Case, Patent Department
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/154,344
FILING DATE: 16-SEP-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/557,146
FILING DATE: 14-DEC-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sterner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-181
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-154-344-2

Query Match 100.0%; Score 40; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9
DB 5 LLLPLQILL 13

RESULT 6
US-08-930-188-2
Sequence 2, Application US/08930188
Patent No. 6093397
GENERAL INFORMATION:
APPLICANT: Dixon, Eric P.
APPLICANT: Johnstone, Edward M.
TITLE OF INVENTION: AMYLOID PRECURSOR PROTEIN PROTEASE AND
TITLE OF INVENTION: RELATED NUCLEIC ACIDS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: United States of America
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/930,188
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/416,257
FILING DATE: 04-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Blalock, Donna K.
REGISTRATION NUMBER: 38,082
REFERENCE/DOCKET NUMBER: X9239
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-277-1090
TELEFAX: 317-276-3861
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-930-188-2

Query Match 100.0%; Score 40; DB 3; Length 253;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9
DB 5 LLLPLQILL 13

RESULT 7
US-09-210-084-3
Sequence 3, Application US/09210084
Patent No. 6197511
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
TITLE OF INVENTION: NOVEL KALLIKREIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/210,084
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/824,874
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0252 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 532504
US-09-210-084-3

Query Match 100.0%; Score 40; DB 3; Length 253;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9
DB 5 LLLPLQILL 13

RESULT 8
US-09-764-762-3
Sequence 3, Application US/09764762
Patent No. 6472195
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
TITLE OF INVENTION: NOVEL KALLIKREIN
NUMBER OF SEQUENCES: 5

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Incyte Pharmaceuticals, Inc.
;; STREET: 3174 Porter Drive
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/764,762
;; FILING DATE: 16-Jan-2001
;; CLASSIFICATION: <Unknown>
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 09/210,084
;; FILING DATE: <Unknown>
;;
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Billings, Lucy J.
;; REGISTRATION NUMBER: 36,749
;; REFERENCE/DOCKET NUMBER: PF-0252 US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-855-0555
;; TELEFAX: 415-845-4166
;;
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 253 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; IMMEDIATE SOURCE:
;; LIBRARY: GenBank
;; CLONE: 532504
;;
;; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-764-762-3

Query Match 100.0%; Score 40; DB 4; Length 253;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQLLL 9
DB 5 LLLPLQLLL 13

RESULT 9
PCT-US96-04294-2
; Sequence 2, Application PC/TUS9604294
; GENERAL INFORMATION:
; APPLICANT: Dixon, Eric P.
; APPLICANT: Johnstone, Edward M.
; APPLICANT: Little, Sheila P.
; TITLE OF INVENTION: AMYLOID PRECURSOR PROTEIN PROTEASE AND
; TITLE OF INVENTION: RELATED NUCLEIC ACIDS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: United States of America
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/04294
; FILING DATE:
; CLASSIFICATION:

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/416,257
;; FILING DATE: 04-APR-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Bialock, Donna K.
;; REGISTRATION NUMBER: 38,082
;; REFERENCE/DOCKET NUMBER: X9239
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 317-277-1090
;; TELEFAX: 317-276-3861
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 253 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; PCT-US96-04294-2

Query Match 100.0%; Score 40; DB 5; Length 253;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQLLL 9
DB 5 LLLPLQLLL 13

RESULT 10
US-09-489-039A-12075
; Sequence 12075, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12075
; LENGTH: 812
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12075

Query Match 95.0%; Score 38; DB 4; Length 812;
Best Local Similarity 88.9%; Pred. No. 21;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQLLL 9
DB 429 LLLPLQLLL 437

RESULT 11
US-09-502-600-35
; Sequence 35, Application US/09502600A
; Patent No. 6294344
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP-C
; CURRENT FILING DATE: 2000-02-11
; CURRENT APPLICATION NUMBER: US/09/502,600A
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 35
; LENGTH: 9

TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Residues 6-14 of the SCCE protein
US-09-502-600-35

Query Match 90.0%; Score 36; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05; 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

QY 2 LLLPLQILL 9
DB 1 LLLPLQILL 8

RESULT 12
US-09-502-600-36
Sequence 36, Application US/09502600A
Patent No. 6294344
GENERAL INFORMATION:
APPLICANT: O'Brien, Timothy J.
TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
TITLE OF INVENTION: Ovarian Cancer
FILE REFERENCE: D6223CIP-C
CURRENT FILING DATE: 2000-02-11
CURRENT APPLICATION NUMBER: US/09/502,600A
PRIOR APPLICATION NUMBER: 09/039,211
PRIOR FILING DATE: 03-14-1998
NUMBER OF SEQ ID NOS: 136
SEQ ID NO 36
LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Residues 4-12 of the SCCE protein
US-09-502-600-36

Query Match 90.0%; Score 36; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQILL 8
DB 2 LLLPLQILL 9

RESULT 13
US-09-918-243-35
Sequence 35, Application US/09918243
Patent No. 6627403
GENERAL INFORMATION:
APPLICANT: O'Brien, Timothy J.
APPLICANT: Cannon, Martin J.
APPLICANT: Santin, Alessandro
TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
FILE REFERENCE: D6223CIP/C/D/CIP
CURRENT APPLICATION NUMBER: US/09/918,243
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US
PRIOR FILING DATE: 2001-07-13
NUMBER OF SEQ ID NOS: 136
SEQ ID NO 35
LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CHAIN
OTHER INFORMATION: Residues 6-14 of the SCCE protein
US-09-918-243-35

Query Match 90.0%; Score 36; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05; 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

QY 2 LLLPLQILL 9
DB 1 LLLPLQILL 8

RESULT 14
US-09-918-243-36
Sequence 36, Application US/09918243
Patent No. 6627403
GENERAL INFORMATION:
APPLICANT: O'Brien, Timothy J.
APPLICANT: Cannon, Martin J.
APPLICANT: Santin, Alessandro
TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
FILE REFERENCE: D6223CIP/C/D/CIP
CURRENT APPLICATION NUMBER: US/09/918,243
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US
PRIOR FILING DATE: 2001-07-13
NUMBER OF SEQ ID NOS: 136
SEQ ID NO 36
LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CHAIN
OTHER INFORMATION: Residues 4-12 of the SCCE protein
US-09-918-243-36

Query Match 90.0%; Score 36; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQILL 8
DB 2 LLLPLQILL 9

RESULT 15
US-08-339-152A-19
Sequence 19, Application US/08339152A
Patent No. 5643726
GENERAL INFORMATION:
APPLICANT: Tanzi, Rudolph E.
APPLICANT: Kovacs, Dora M.
TITLE OF INVENTION: Methods For Modulating Transcription
TITLE OF INVENTION: From The Amyloid -Protein Precursor (APP) Promoter
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 New York Ave., NW, Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/339,152A
FILING DATE: 10-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Steffe, Eric K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 0609.4120000
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
TELEX:

; INFORMATION FOR SEQ ID NO: 19:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 190 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

US-08-339-152A-19

Query Match 82.5%; Score 33; DB 1; Length 190;

Best Local Similarity 77.8%; Pred. No. 41;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9

DB 1 LLLPLSLLL 9

Search completed: March 1, 2004, 17:38:23

Job time : 11.8889 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:16:55 ; Search time 45 5556 Seconds

(without alignments)
55.820 Million cell updates/sec

Title: US-09-905-083-33

Perfect score: 40

Sequence: 1 LLLPLQLL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	40	100.0	9	AAE08238	Aae08238 Human str
2	40	100.0	136	ABG23378	Abg23378 Novel hum
3	40	100.0	198	ADA05736	Ada05736 Human NOV
4	40	100.0	250	ADA05732	Ada05732 Human NOV
5	40	100.0	253	AAE07888	Aae07888 Human str
6	40	100.0	253	AAW05383	Aaw05383 Human any
7	40	100.0	253	ABB84421	Abb84421 Human SCC
8	40	100.0	253	ABB84406	Abb84406 Human SCC
9	40	100.0	253	AAU02740	Aau02740 Amino aci
10	40	100.0	253	ABU07440	Abu07440 Protein d
11	40	100.0	253	ABR58471	Abu07471 Protein d
12	40	100.0	253	ABR58471	Abu07471 Human str
13	40	100.0	253	ADB80484	Adb80484 Ovarian c
14	40	100.0	257	ADA05732	Ada05732 Human HSC
15	36	90.0	9	AAE08240	Aae08240 Human str
16	36	90.0	9	AAE08241	Aae08241 Human str
17	35	87.5	61	AAO12472	Aao12472 Human pol
18	35	87.5	142	ABG63580	Abg63580 Human gas
19	35	87.5	156	ABG63578	Abg63578 Human gas
20	35	87.5	159	ABG63582	Abg63582 Human gas
21	34	85.0	23	ABG43858	Abg43858 Peptide #
22	34	85.0	23	AAE08241	Aae08241 Peptide #
23	34	85.0	23	AAE08241	Aae08241 Peptide #
24	34	85.0	23	AAE08241	Aae08241 Peptide #
25	34	85.0	23	ABG46617	Abg46617 Human pep

26	34	85.0	201	5	ABB78636	Abb78636 Rat OST10
27	34	85.0	371	2	AAR75642	Aar75642 Bovine co
28	34	85.0	778	6	ABM68947	Abm68947 Phototrab
29	34	85.0	845	6	ABU23951	Abu23951 Protein e
30	33	82.5	12	5	ABG98371	Abg98371 Secreted
31	33	82.5	152	6	ABP75437	Abp75437 Human sec
32	33	82.5	176	5	AAE22748	Aae22748 Consensus
33	33	82.5	190	2	AAR98922	Aar98922 Murine AP
34	33	82.5	562	5	ABB09701	Abb09701 Amino aci
35	33	82.5	653	2	AAR98903	Aar98903 Murine AP
36	33	82.5	770	4	ABG25713	Abg25713 Novel hum
37	33	82.5	1404	4	ABG14929	Abg14929 Novel hum
38	32	80.0	21	2	AAW48387	Aaw48387 Synthetic
39	32	80.0	21	6	ABG71643	Abg71643 Leader se
40	32	80.0	22	2	AAW59850	Aaw59850 Amino aci
41	32	80.0	22	5	AAU10579	Aau10579 Mammalian
42	32	80.0	71	4	AAW89706	Aaw89706 Human imm
43	32	80.0	73	4	ABG19347	Abg19347 Novel hum
44	32	80.0	102	7	ADB64112	Adb64112 Human pro
45	32	80.0	147	2	AAV59764	Aay59764 Human nor

ALIGNMENTS

RESULT 1

AAE08238
ID AAE08238 standard; peptide; 9 AA.

XX AAE08238;

XX 01-NOV-2001 (first entry)

XX Human stratum corneum chymotrypsin enzyme peptide #3 (residues 5-13).

XX Stratum corneum chymotrypsin enzyme; SCCE; cytostatic; vaccine; tumour;
cancer; ovarian; breast; lung; colon; prostate; carcinoma; human;
antitense therapy; malignant hyperplasia.

XX Homo sapiens.

XX WO200159158-A1.

XX 16-AUG-2001.

XX 07-FEB-2001; 2001WO-US003977.

XX 11-FEB-2000; 2000US-00502600.

XX (UVAR-) UNIV ARKANSAS.

XX O'brien TU;

XX WPI; 2001-514676/56.

XX Diagnosing cancer comprises detecting stratum corneum chymotrypsin enzyme.

XX Claim 25; Page 102; 127pp; English.

XX The invention relates to diagnosing cancer especially ovarian cancer, by screening for stratum corneum chymotrypsin enzyme (SCCE). Proteases are considered to be an integral part of tumour growth and metastasis, and therefore, markers indicative of their presence or absence are useful for the diagnosis of cancer. The method is useful for diagnosing cancer, carcinoma and malignant hyperplasia. The SCCE oligonucleotide may be used to treat a cancer selected from ovarian, breast, lung, colon, prostate and other cancers in which SCCE is overexpressed. The present sequence is human SCCE peptide

XX Sequence 9 AA;

XX Query Match 100.0%; Score 40; DB 4; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPQLILL 9
| | | | |
Db 1 LLLPQLILL 9

RESULT 2
ABG23378
ID ABG23378 standard; protein; 136 AA.

XX AC ABG23378;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #23369.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.

XX PN WO200175067-A2.
XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.
XX PI Dmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.

XX DR N-PSDB; AAS87565.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX PS Claim 20; SEQ ID NO 53737; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activities. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 136 AA;

Query Match 100.0%; Score 40; DB 4; Length 136;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPQLILL 9
| | | | |
Db 11 LLLPQLILL 19

RESULT 3
ADA05736
ID ADA05736 standard; protein; 198 AA.

XX AC ADA05736;
XX DT 06-NOV-2003 (first entry)
XX DE Human NOV18c protein SEQ ID NO:96.

XX KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
KW immunomodulator; cytostatic; nootropic; neuroprotective;
KW antiparkinsonian; antilipemic; gene therapy; human disease;
KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW immune disorder; haematopoietic disorder; dyslipidaemia.

XX CS Homo sapiens.

XX PN WO2003029424-A2.

XX PD 10-APR-2003.

XX PF 02-OCT-2002; 2002WO-US013173.

XX PR 02-OCT-2001; 2001US-0326483P.

XX PR 05-OCT-2001; 2001US-0327435P.

XX PR 09-OCT-2001; 2001US-0327917P.

XX PR 09-OCT-2001; 2001US-0328029P.

XX PR 09-OCT-2001; 2001US-0328044P.

XX PR 09-OCT-2001; 2001US-0328056P.

XX PR 12-OCT-2001; 2001US-0328849P.

XX PR 15-OCT-2001; 2001US-0328941P.

XX PR 17-OCT-2001; 2001US-0330142P.

XX PR 18-OCT-2001; 2001US-0330309P.

XX PR 22-OCT-2001; 2001US-0341059P.

XX PR 24-OCT-2001; 2001US-0339266P.

XX PR 24-OCT-2001; 2001US-0343639P.

XX PR 29-OCT-2001; 2001US-0349575P.

XX PR 01-NOV-2001; 2001US-0346357P.

XX PR 17-APR-2002; 2002US-0373260P.

XX PR 19-APR-2002; 2002US-0373815P.

XX PR 19-APR-2002; 2002US-0373817P.

XX PR 19-APR-2002; 2002US-0373826P.

XX PR 19-APR-2002; 2002US-0373884P.

XX PR 22-APR-2002; 2002US-0374977P.

XX PR 16-MAY-2002; 2002US-0381037P.

XX PR 16-MAY-2002; 2002US-0381038P.

XX PR 17-MAY-2002; 2002US-0381042P.

XX PR 26-MAY-2002; 2002US-0381642P.

XX PR 29-MAY-2002; 2002US-0383656P.

XX PR 25-JUN-2002; 2002US-0383831P.

XX PR 01-OCT-2002; 2002US-03862511.

(CURA-) CURAGEN CORP.

XX PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;

XX PI Paturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UN;

XX PI Ort T, Gorman L, Zertusen BD, Anderson DW, Zhong M, Catterton E;

XX PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shency SG;

XX PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;

XX PI Eissen AJ, Gangolli EA, Rieger DK, Spaderna SK;

XX WPI; 2003-381626/36.

XX DR N-PSDB; ADA05735.

XX New NOVX polypeptides and nucleic acids, useful for diagnosing,
PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
XX Claim 1; Page 170; 586pp; English.
XX
XX The present invention describes NOVX proteins, where X can be 1 to 55
CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide
CC described above and a carrier; (2) a kit comprising, in one or more
CC containers, the composition described above; (3) an isolated nucleic acid
CC molecule which encodes a NOVX protein of the invention; (4) a vector
CC comprising the nucleic acid molecule described above; (5) a cell
CC comprising the above vector; (6) an antibody that immunospecifically
CC binds to the polypeptide described above; (7) methods for determining the
CC presence or amount of the above polypeptide or nucleic acid molecule in a
CC sample; (8) methods for determining the presence of or predisposition to
CC a disease associated with altered levels of expression of the above
CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
CC method of identifying an agent that binds to the polypeptide described
CC above; (10) a method for identifying a potential therapeutic agent for
CC use in treating a pathology that is related to an aberrant expression or
CC aberrant physiological interactions of the polypeptide; (11) a method of
CC screening for a modulator of activity or of latency or predisposition to
CC a pathology associated with the polypeptide; (12) a method for modulating
CC the activity of the polypeptide described above; (13) methods of treating
CC or preventing a pathology associated with the above polypeptide in a
CC mammal; and (14) a method for producing the above polypeptide. NOVX
CC sequences have antidiabetic, anorectic, antibacterial, virucide,
CC immunomodulator, cytostatic, nontropic, neuroprotective, antiparkinsonian
CC and antilipidemic activities, and can be used in gene therapy. The
CC polypeptide is useful in manufacturing a medicament for treating a
CC syndrome associated with a human disease. The polypeptide or the nucleic
CC acid molecule may be used to diagnose, treat or prevent metabolic
CC disorders such as diabetes or obesity, infections, cachexia, cancer,
CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
CC disease, immune disorders, haematopoietic disorders and various
CC dyslipidaemias. The nucleic acids can also be used as hybridisation
CC probes, in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. The present sequence represents a human NOVX from the
CC present invention.
XX
XX Sequence 198 AA;
XX
XX Query Match 100.0%; Score 40; DB 6; Length 198;
XX Best Local Similarity 100.0%; Pred. No. 13;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 LLLPLQLLL 9
XX Db 5 LLLPLQLLL 13
XX
XX RESULT 4
XX ADA05732
XX ID ADA05732 standard; protein; 250 AA.
XX AC ADA05732;
XX XX
XX DT 06-NOV-2003 (first entry)
XX XX
XX DE Human NOV18a protein SEQ ID NO:92.
XX KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
XX KW immunomodulator; cytostatic; nontropic; neuroprotective;
XX KW antiparkinsonian; antilipidemic; gene therapy; human disease;
XX KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
XX KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX KW immune disorder; haematopoietic disorder; dyslipidaemia.
XX OS Homo sapiens.
XX

PN WO2003029424-A2.
XX
XX 10-APR-2003.
XX
XX 02-OCT-2002; 2002WO-US031373.
XX
XX 02-OCT-2001; 2001US-0326483P.
XX PR 05-OCT-2001; 2001US-0327435P.
XX PR 09-OCT-2001; 2001US-0327449P.
XX PR 09-OCT-2001; 2001US-0327917P.
XX PR 09-OCT-2001; 2001US-0328029P.
XX PR 09-OCT-2001; 2001US-0328044P.
XX PR 09-OCT-2001; 2001US-0328056P.
XX PR 12-OCT-2001; 2001US-0328849P.
XX PR 15-OCT-2001; 2001US-0329414P.
XX PR 17-OCT-2001; 2001US-0330142P.
XX PR 18-OCT-2001; 2001US-0330309P.
XX PR 22-OCT-2001; 2001US-0341058P.
XX PR 24-OCT-2001; 2001US-0339256P.
XX PR 24-OCT-2001; 2001US-0343629P.
XX PR 29-OCT-2001; 2001US-0349575P.
XX PR 01-NOV-2001; 2001US-0346357P.
XX PR 17-APR-2002; 2002US-0373260P.
XX PR 19-APR-2002; 2002US-0373815P.
XX PR 19-APR-2002; 2002US-0373817P.
XX PR 19-APR-2002; 2002US-0373826P.
XX PR 19-APR-2002; 2002US-0373884P.
XX PR 22-APR-2002; 2002US-0374977P.
XX PR 16-MAY-2002; 2002US-0381037P.
XX PR 16-MAY-2002; 2002US-0381038P.
XX PR 17-MAY-2002; 2002US-0381042P.
XX PR 17-MAY-2002; 2002US-0381642P.
XX PR 28-MAY-2002; 2002US-0383556P.
XX PR 29-MAY-2002; 2002US-0383831P.
XX PR 25-JUN-2002; 2002US-0391335P.
XX PR 01-OCT-2002; 2002US-00262511.
XX (CURA-) CURAGEN CORP.
XX
XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
XX Faturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
XX Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;
XX Ji W, Miller CE, Rastelli L, Stone DU, Pena CE, Shenoy SC;
XX Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;
XX Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;
XX
XX WPI: 2003-381626/36.
XX DR N-PSDB; ADA05731.
XX
XX New NOVX polypeptides and nucleic acids, useful for diagnosing,
PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
PT
PT
PT
XX
XX Claim 1; Page 169-170; 586pp; English.
XX
XX The present invention describes NOVX proteins, where X can be 1 to 55
CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide
CC described above and a carrier; (2) a kit comprising, in one or more
CC containers, the composition described above; (3) an isolated nucleic acid
CC molecule which encodes a NOVX protein of the invention; (4) a vector
CC comprising the nucleic acid molecule described above; (5) a cell
CC comprising the above vector; (6) an antibody that immunospecifically
CC binds to the polypeptide described above; (7) methods for determining the
CC presence or amount of the above polypeptide or nucleic acid molecule in a
CC sample; (8) methods for determining the presence of or predisposition to
CC a disease associated with altered levels of expression of the above
CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
CC polypeptide or nucleic acid molecule that binds to the polypeptide described
CC above; (10) a method for identifying a potential therapeutic agent for
CC use in treating a pathology that is related to an aberrant expression or
CC aberrant physiological interactions of the polypeptide; (11) a method of
CC screening for a modulator of activity or of latency or predisposition to

CC a pathology associated with the polypeptide; (12) a method for modulating
 CC the activity of the polypeptide described above; (13) methods of treating
 CC or preventing a pathology associated with the above polypeptide in a
 CC mammal; and (14) a method for producing the above polypeptide. NOVX
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,
 CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian
 CC and antilipemic activities, and can be used in gene therapy. The
 CC polypeptide is useful in manufacturing a medicament for treating a
 CC syndrome associated with a human disease. The polypeptide or the nucleic
 CC acid molecule may be used to diagnose, treat or prevent metabolic
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
 CC disease, immune disorders, haematopoietic disorders and various
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation
 CC probes, in chromosome mapping, tissue typing, preventive medicine and
 CC pharmacogenomics. The present sequence represents a human NOVX from the
 CC present invention.

XX SQ Sequence 250 AA;

Query Match 100.0%; Score 40; DB 6; Length 250;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQLLL 9
 Db 2 LLLPLQLLL 10

RESULT 5

AA067888
 ID AAR67888 standard; protein; 253 AA.

XX AC AAR67888;

XX DT 25-MAR-2003 (revised)

XX DT 09-AUG-1995 (first entry)

XX DE Human stratum corneum chymotryptic recombinant enzyme (SCCE).

XX KW Stratum corneum chymotryptic enzyme; skin disorder; acne; psoriasis;
 XX KW callosities; keratosis pilaris; ichthyoses; eczema.

XX OS Homo sapiens.

XX PN WO9500651-A1.

XX PD 05-JAN-1995.

XX PF 20-JUN-1994; 94WO-IB000166.

XX PR 18-JUN-1993; 93DK-00000725.

XX PA (SYMB-) SYMBICOM AB.

XX PI Egelrud T, Hansson L;

XX DR WPI; 1995-052088/07.

XX DR N-PSDB; AAO81203.

XX PT Nucleotide sequences encoding stratum corneum chymotryptic enzyme - and
 XX PT related vectors, transformed cells and polypeptides, useful for treating
 XX PT skin disorders, e.g. acne or psoriasis, and for identification of
 XX PT specific inhibitors.

XX PS Disclosure; Page 97; 137pp; English.

XX CC The enzyme encoded by this sequence is used in pharmaceutical, cosmetic
 XX CC and skin care products, especially to treat and prevent acne, xeroderma,
 XX CC or other hyperkeratotic conditions (e.g. callosities or keratosis
 XX CC pilaris), ichthyoses, psoriasis, eczema, etc. It is produced
 XX CC recombinantly following mammal, insect, plant, or microorganism
 XX CC transformation with plasmid pS507. (Updated on 25-MAR-2003 to correct PN

CC field.)

XX SQ Sequence 253 AA;

Query Match 100.0%; Score 40; DB 2; Length 253;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQLLL 9
 Db 5 LLLPLQLLL 13

RESULT 6

AA05383

ID AAW05383 standard; protein; 253 AA.

XX AC AAW05383;

XX DT 31-DEC-1996 (first entry)

XX DE Human amyloid precursor protein protease.

XX KW Amyloid precursor protein protease; Alzheimer's disease; diagnosis;
 XX KW therapy.

XX OS Homo sapiens.

XX PN WO9631122-A1.

XX PD 10-OCT-1996.

XX PF 02-APR-1996; 96WO-US004294.

XX PR 04-APR-1995; 95US-00416257.

XX PA (ELIL) LILLY & CO ELI.

XX PI Dixon EP, Johnstone EM, Little SP;

XX DR WPI; 1996-464694/46.

XX DR N-PSDB; AAT39783.

XX PT New isolated human amyloid precursor protein protease - used to develop
 XX PT prods. for the treatment or diagnosis of associated conditions, esp.
 XX PT Alzheimer's disease.

XX PS Claim 1; Page 44-45; 55pp; English.

XX CC Human amyloid precursor protein protease (AAW05383) is involved in the
 XX CC processing or clearance of amyloid precursor protein to form beta-amyloid
 XX CC peptide. Its amino acid sequence was deduced from a cDNA clone (AAT39783)
 XX CC obt'd. from a human lung library. Recombinant protease can be produced in
 XX CC transformed or transfected prokaryotic (partic. E. coli) or eukaryotic
 XX CC (partic. AV-120 host cells. It is used to develop products for the design
 XX CC and testing of cpds. useful for treating or preventing conditions
 XX CC associated with beta-amyloid peptide, esp. Alzheimer's disease

XX SQ Sequence 253 AA;

Query Match 100.0%; Score 40; DB 2; Length 253;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQLLL 9
 Db 5 LLLPLQLLL 13

RESULT 7

ABB84421

ID ABB84421 standard; peptide; 253 AA.

XX

AC ABB84421;
 XX 08-NOV-2002 (first entry)
 XX Human SCCE protein N-terminal fragment SEQ ID 48.
 XX
 XX SCCE; human; stratum corneum chymotryptic enzyme; kallikrein 7;
 KW serine protease; transgenic mammal; skin; skin disease; skin cancer;
 KW hyperkeratosis; acanthosis; epidermal inflammation; dermal inflammation;
 XX pruritus; atopic dermatitis; eczema; acne; itch; KLK7; enzyme.
 OS Homo sapiens.
 XX WO200262135-A2.
 XX 15-AUG-2002.
 XX
 XX 08-FEB-2002; 2002WO-IB001300.
 XX 09-FEB-2001; 2001CA-02332655.
 XX 09-FEB-2001; 2001DK-00000218.
 XX (EGEL/) EGELRUD T.
 XX (HANS/) HANSSON L.
 XX
 XX Egelrud T, Hansson L;
 XX WPI; 2002-643380/69.
 XX Transgenic mammal or its embryo useful as model for human disease, has
 PT heterologous nucleotide sequence coding for stratum corneum chymotryptic
 PT enzyme operably linked to promoter that drives its expression in skin.
 XX
 XX Example 6; Page 37; 74pp; English.
 XX This invention describes a novel non-human transgenic mammal or mammalian
 CC embryo having integrated within its genome, a heterologous nucleotide
 CC sequence comprising at least a significant part of a nucleotide sequence
 CC coding for a stratum corneum chymotryptic enzyme (SCCE) or its variant,
 CC operably linked to a promoter that drives expression of heterologous scce
 CC or its variant in skin. The product of the invention is useful as a model
 CC for the study of disease with the aim of improving treatment, to relieve
 CC or ameliorate a pathological condition, for development or testing of a
 CC cosmetic or a pharmaceutical formulation, and for the development of a
 CC diagnostic method. It can also be used as a model for a skin disease or
 CC skin cancer. The invention is also useful for screening or identifying a
 CC compound or composition effective for the prevention or treatment of an
 CC abnormal or unwanted phenotype, and for screening or identifying a
 CC inflammatory skin diseases selected from diseases consisting of epidermal
 CC hyperkeratosis, acanthosis, epidermal inflammation, dermal inflammation,
 CC pruritus, atopic dermatitis, eczema, acne and inherited skin diseases
 CC with epidermal hyperkeratosis. The mammal of the invention is also useful
 CC as a model for further studies of itch mechanisms and the testing of
 CC potential compounds and compositions for relieve of various skin diseases
 CC where itch is a component. This sequence represents the N-terminal
 CC fragment of the human stratum corneum chymotryptic enzyme, SCCE
 CC synonymous with human kallikrein 7 (KLK7), used in the development of the
 CC transgenic mammals described in the invention
 XX
 XX Sequence 253 AA;

Query Match 100.0%; Score 40; DB 5; Length 253;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9
 |||||
 Db 5 LLLPLQILL 13

RESULT 8
 ABB84406

ID ABB84406 standard; protein; 253 AA.
 XX ABB84406;
 XX 08-NOV-2002 (first entry)
 XX Human SCCE protein.
 XX
 XX SCCE; human; stratum corneum chymotryptic enzyme; kallikrein 7;
 KW serine protease; transgenic mammal; skin; skin disease; skin cancer;
 KW hyperkeratosis; acanthosis; epidermal inflammation; dermal inflammation;
 KW pruritus; atopic dermatitis; eczema; acne; itch; KLK7; enzyme.
 XX Homo sapiens.
 XX WO200262135-A2.
 XX 15-AUG-2002.
 XX
 XX 08-FEB-2002; 2002WO-IB001300.
 XX 09-FEB-2001; 2001CA-02332655.
 XX 09-FEB-2001; 2001DK-00000218.
 XX (EGEL/) EGELRUD T.
 XX (HANS/) HANSSON L.
 XX
 XX Egelrud T, Hansson L;
 XX WPI; 2002-643380/69.
 XX N-PSDB; ABQ76226.
 XX Transgenic mammal or its embryo useful as model for human disease, has
 PT heterologous nucleotide sequence coding for stratum corneum chymotryptic
 PT enzyme operably linked to promoter that drives its expression in skin.
 XX
 XX Claim 10; Page 58-59; 74pp; English.
 XX This invention describes a novel non-human transgenic mammal or mammalian
 CC embryo having integrated within its genome, a heterologous nucleotide
 CC sequence comprising at least a significant part of a nucleotide sequence
 CC coding for a stratum corneum chymotryptic enzyme (SCCE) or its variant,
 CC operably linked to a promoter that drives expression of heterologous scce
 CC or its variant in skin. The product of the invention is useful as a model
 CC for the study of disease with the aim of improving treatment, to relieve
 CC or ameliorate a pathological condition, for development or testing of a
 CC cosmetic or a pharmaceutical formulation, and for the development of a
 CC diagnostic method. It can also be used as a model for a skin disease or
 CC skin cancer. The invention is also useful for screening or identifying a
 CC compound or composition effective for the prevention or treatment of an
 CC abnormal or unwanted phenotype, and for screening or identifying a
 CC inflammatory skin diseases selected from diseases consisting of epidermal
 CC hyperkeratosis, acanthosis, epidermal inflammation, dermal inflammation,
 CC pruritus, atopic dermatitis, eczema, acne and inherited skin diseases
 CC with epidermal hyperkeratosis. The mammal of the invention is also useful
 CC as a model for further studies of itch mechanisms and the testing of
 CC potential compounds and compositions for relieve of various skin diseases
 CC where itch is a component. This sequence represents the human stratum
 CC corneum chymotryptic enzyme, SCCE which is a serine protease synonymous
 CC with human kallikrein 7 (KLK7) and is used in the development of the
 CC transgenic mammals described in the invention
 XX
 XX Sequence 253 AA;

Query Match 100.0%; Score 40; DB 5; Length 253;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9
 |||||
 Db 5 LLLPLQILL 13

Query Match 100.0%; Score 40; DB 6; Length 253;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQILL 9
Db 5 LLLPLQILL 13

RESULT 11

ID ABU07471 standard; protein; 253 AA.

XX ABU07471;

AC

DT 28-JAN-2003 (first entry)

DE Protein differentially regulated in prostate cancer #74.

DE

DE

KW Prostate cancer; gene expression; differential regulation;
molecular marker; drug target; cancer detection; cancer diagnosis;
cancer staging; cancer grading; cancer assessing; cancer monitoring.

KW

XX Homo sapiens.

OS

XX WO200281638-A2.

PN

XX

PD 17-OCT-2002.

PD

XX

XX 06-APR-2002; 2002WO-US010824.

XX

XX

PR 06-APR-2001; 2001US-0281731P.

PR

XX

XX 06-APR-2001; 2001US-0281732P.

XX

XX (ORIG-) ORIGENE TECHNOLOGIES INC.

PA

XX

XX

PI Sun Z, Jay G;

XX

XX WPI; 2003-058520/05.

DR

DR N-PSDB; ABX10375.

XX

XX

PT Novel genes which are differentially regulated in prostate cancer, useful
for diagnosing prostate cancer in prostate tissue sample and assessing
therapeutic or preventive intervention in prostate cancer patients.

PT

XX

XX Claim 1; Page 351; 416pp; English.

XX

XX The invention describes genes (I) which are differentially regulated in
prostate cancer. (I) Is useful for diagnosing a prostate cancer in a
sample comprising prostate tissue, which involves determining the number
of target genes which are differentially-regulated in the sample, where
the number is indicative of the probability that the sample comprises
prostate cancer. (I) Is useful for assessing a therapeutic or preventive
intervention in a subject having a prostate cancer, which involves
determining the expression levels in a sample comprising prostate tissue
of target genes which are differentially-regulated in prostate cancer.
Preferably, the expression levels of at least 10 genes are determined.
(I) is also useful for identifying agents that modulate a biological
activity of a polypeptide differentially-regulated in prostate cancer
cells, which involves contacting a polypeptide differentially-regulated
in prostate cancer cells with a test agent under conditions effective for
the test agent to modulate a biological activity of the polypeptide, and
determining whether the test agent modulates the biological activity. (I)
is useful as molecular markers, as drug targets, and for detecting,
diagnosing, staging, grading, assessing, monitoring, prognosticating,
preventing or treating, determining predisposition to diseases and
conditions especially relating to prostate cancer. (I) and its expression
products are used in the diagnostic test to assay for presence of cancer
e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
blood etc. (I) is useful for assessing cancer e.g., to determine the type
of cancer, its stage of development, the nature of genetic defect, etc.
The polypeptide encoded by (I) can be used as target for therapy or drug
discovery. (I) can also be used for expressing the polypeptide and thus

CC for searching specific binding partners of the polypeptide. (I) is useful
in therapeutic applications to treat prostate cancer. The identification
of specific genes, and groups of genes, expressed in pathways
physiologically relevant to prostate cancer permits the definition of
functional and disease pathways and the delineation of targets in these
pathways which are useful in diagnostic, therapeutic, and clinical
applications. This is the amino acid sequence of a protein differentially
regulated in prostate cancer

XX

SQ Sequence 253 AA;

Query Match 100.0%; Score 40; DB 6; Length 253;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQILL 9
Db 5 LLLPLQILL 13

RESULT 12

ABR58471

ID ABR58471 standard; protein; 253 AA.

XX ABR58471;

AC

DT 07-JUL-2003 (first entry)

DE Human stratum corneum chymotryptic enzyme - ovarian cancer clone O1676P.

DE

DE

KW Human; cytostatic; gene therapy; vaccine; cancer; ovarian cancer.

KW

XX Homo sapiens.

OS

XX WO2003029468-A1.

PN

XX

PD 10-APR-2003.

PD

XX

XX 02-OCT-2002; 2002WO-US031467.

XX

PF

XX

PR 02-OCT-2002; 2001US-0327135P.

PR

XX

XX 30-MAY-2002; 2002US-0384531P.

XX

XX (CORI-) CORIXA CORP.

PA

XX

XX Algate PA, Mannion J;

PI

XX WPI; 2003-372001/35.

DR

XX

XX New polynucleotide and polypeptide useful for diagnosing and/or treating
cancer, particularly ovarian cancer, and as a vaccine.

PT

XX

XX Claim 2; Page 157-158; 169pp; English.

XX

XX The invention relates to a novel isolated polynucleotide. The
polynucleotides of the invention have cytostatic activity, and may have a
use in gene therapy, and in a vaccine. The composition and methods are
useful in diagnosing and/or treating cancer, particularly ovarian cancer.
The composition may also be used as a vaccine to prevent cancer. The
present sequence is used in the exemplification of the invention

XX

SQ Sequence 253 AA;

Query Match 100.0%; Score 40; DB 6; Length 253;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQILL 9
Db 5 LLLPLQILL 13

RESULT 13

PF 07-FEB-2001; 2001WO-US003977.
XX
PR 11-FEB-2000; 2000US-00502600.
XX
PA (UYAR-) UNIV ARKANSAS.
XX
XX O'brien TJ;
PI
XX
DR WPI; 2001-S14676/56.
XX
XX
PT Diagnosing cancer comprises detecting stratum corneum chymotrypsin
PT enzyme.
XX
XX
PS Claim 25; Page 103; 127pp; English.
XX
XX The invention relates to diagnosing cancer especially ovarian cancer, by
CC screening for stratum corneum chymotrypsin enzyme (SCCE). Proteases are
CC considered to be an integral part of tumour growth and metastasis, and
CC therefore, markers indicative of their presence or absence are useful for
CC the diagnosis of cancer. The method is useful for diagnosing cancer,
CC carcinoma and malignant hyperplasia. The SCCE oligonucleotide may be used
CC to treat a cancer selected from ovarian, breast, lung, colon, prostate
CC and other cancers in which SCCE is overexpressed. The present sequence is
CC human SCCE peptide
XX
SQ Sequence 9 AA;
Query Match 90.0%; Score 36; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. NO. 1.4e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Caps 0;
Qy 2 LLPLQILL 9
Db 1 LLPLQILL 8
Search completed: March 1, 2004, 17:28:49
Job time : 47.5556 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:35:01 ; Search time 24.1111 Seconds
(without alignments)
78.818 Million cell updates/sec

Title: US-09-905-083-33
Perfect score: 40
Sequence: 1 LLLPLQILL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/prodata/1/pubaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/1/pubaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata/1/pubaa/US05_NEW_PUB.pep.*
4: /cgn2_6/prodata/1/pubaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/1/pubaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/prodata/1/pubaa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata/1/pubaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/1/pubaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/prodata/1/pubaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/prodata/1/pubaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/prodata/1/pubaa/US09_NEW_PUB.pep.*
13: /cgn2_6/prodata/1/pubaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/prodata/1/pubaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/prodata/1/pubaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/prodata/1/pubaa/US10_NEW_PUB.pep.*
17: /cgn2_6/prodata/1/pubaa/US60_NEW_PUB.pep.*
18: /cgn2_6/prodata/1/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	9	9	US-09-918-243-33
2	40	100.0	9	9	US-09-905-083-33
3	40	100.0	253	9	US-09-888-615-98
4	40	100.0	253	9	US-09-764-762-3
5	40	100.0	253	14	US-10-264-283-90
6	40	100.0	253	15	US-10-295-027-498
7	40	100.0	253	15	US-10-173-999-48
8	36	90.0	9	9	US-09-918-243-35
9	36	90.0	9	9	US-09-918-243-36
10	36	90.0	9	9	US-09-905-083-35
11	36	90.0	9	9	US-09-905-083-36
12	34	85.0	23	9	US-09-864-761-46097
13	34	85.0	201	10	US-09-956-622A-39
14	33	82.5	516	14	US-10-156-761-11235
15	32	80.0	21	13	US-10-096-241-24

16	32	80.0	102	15	US-10-104-047-2266
17	32	80.0	218	14	US-10-378-393-20
18	32	80.0	219	14	US-10-378-393-3
19	32	80.0	219	14	US-10-378-393-13
20	32	80.0	235	14	US-10-207-655-174
21	32	80.0	302	9	US-09-393-634-37
22	32	80.0	302	10	US-09-510-332-3
23	32	80.0	302	14	US-10-383-983-37
24	32	80.0	418	15	US-10-295-027-181
25	32	80.0	500	14	US-10-156-761-15018
26	32	80.0	559	14	US-10-156-761-15018
27	32	80.0	574	10	US-09-374-046A-62
28	32	80.0	624	14	US-10-378-393-7
29	32	80.0	646	14	US-10-378-393-18
30	32	80.0	791	14	US-10-378-393-11
31	32	80.0	868	15	US-10-369-493-10255
32	32	80.0	1306	14	US-10-017-724-4
33	31	77.5	34	9	US-09-864-761-39601
34	31	77.5	40	9	US-09-864-761-39134
35	31	77.5	46	14	US-10-050-882-135
36	31	77.5	92	11	US-09-833-245-1829
37	31	77.5	92	11	US-09-833-245-1830
38	31	77.5	92	11	US-09-833-245-1831
39	31	77.5	99	11	US-09-833-245-1146
40	31	77.5	103	15	US-10-264-237-1142
41	31	77.5	110	9	US-09-948-018-17
42	31	77.5	114	14	US-10-106-698-4375
43	31	77.5	120	9	US-09-764-877-1192
44	31	77.5	120	15	US-10-242-515-1192
45	31	77.5	146	11	US-09-833-245-2081

ALIGNMENTS

RESULT 1

US-09-918-243-33
; Sequence 33, Application US/09918243
; Patent No. US20020142317A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 33
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 5-13 of the SCCE protein
US-09-918-243-33

Query Match 100.0%; Score 40; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQILL 9

Db 1 LLLPLQILL 9

RESULT 2

US-09-905-083-33
; Sequence 33, Application US/09905083
; Patent No. US20020146708A1
; GENERAL INFORMATION:

APPLICANT: O'Brien, Timothy J.
TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
FILE OF INVENTION: Ovarian Cancer
FILE REFERENCE: D6223CIP/C/DIV
CURRENT APPLICATION NUMBER: US/09/905,083
CURRENT FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: US 09/502,600
PRIOR FILING DATE: 2000-02-11
NUMBER OF SEQ ID NOS: 136
SEQ ID NO 33
LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CHAIN
OTHER INFORMATION: Residues 5-13 of the SCCE protein
US-09-905-083-33

Query Match 100.0%; Score 40; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQLLL 9
Db 1 LLLPLQLLL 9

RESULT 3
US-09-888-615-98

Sequence 98, Application US/09888615
Patent No. US20020064856A1
GENERAL INFORMATION:
APPLICANT: PLOWMAN, GREGORY
APPLICANT: WHYTE, DAVID
APPLICANT: CAENEPEEL, SEAN
APPLICANT: CHARYDCZAK, GLEN
APPLICANT: MANNING, GERARD
APPLICANT: SUDARSANAM, SUCHA
TITLE OF INVENTION: NOVEL PROTEASES
FILE REFERENCE: 038602/1214
CURRENT APPLICATION NUMBER: US/09/888,615
CURRENT FILING DATE: 2001-06-26
PRIOR APPLICATION NUMBER: 60/214,047
PRIOR FILING DATE: 2000-06-26
NUMBER OF SEQ ID NOS: 150
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 98
LENGTH: 253
TYPE: PRT
ORGANISM: Homo sapiens
US-09-888-615-98

Query Match 100.0%; Score 40; DB 9; Length 253;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQLLL 9
Db 5 LLLPLQLLL 13

RESULT 4
US-09-764-762-3

Sequence 3, Application US/09764762
Patent No. US20020068341A1
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: NOVEL KALLIKREIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive

CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/764,762
FILING DATE: 16-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/210,084
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0252 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 532504
SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-764-762-3

Query Match 100.0%; Score 40; DB 9; Length 253;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQLLL 9
Db 5 LLLPLQLLL 13

RESULT 5
US-10-264-283-90

Sequence 90, Application US/10264283
Publication No. US20030144494A1
GENERAL INFORMATION:
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
FILE REFERENCE: 210121.590
CURRENT APPLICATION NUMBER: US/10/264,283
CURRENT FILING DATE: 2002-10-02
NUMBER OF SEQ ID NOS: 111
SOFTWARE: Corixa Invention Disclosure Database
SEQ ID NO 90
LENGTH: 253
TYPE: PRT
ORGANISM: Homo sapiens
US-10-264-283-90

Query Match 100.0%; Score 40; DB 14; Length 253;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQLLL 9
Db 5 LLLPLQLLL 13

```

RESULT 6
US-10-295-027-498
; Sequence 498, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afari, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 498
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-498

Query Match      100.0%; Score 40; DB 15; Length 253;
Best Local Similarity 100.0%; Pred. No. 19;
Matches      9; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      1 LLLPLQILL 9
Db      5 LLLPLQILL 13

RESULT 7
US-10-173-999-48
; Sequence 48, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002420US
; CURRENT APPLICATION NUMBER: US/10/173,999
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234

```

```

; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-173-999-48

Query Match      100.0%; Score 40; DB 15; Length 253;
Best Local Similarity 100.0%; Pred. No. 19;
Matches      9; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      1 LLLPLQILL 9
Db      5 LLLPLQILL 13

RESULT 8
US-09-918-243-35
; Sequence 35, Application US/09918243
; Patent No. US20020142317A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 35
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 6-14 of the SCCE protein
US-09-918-243-35

Query Match      90.0%; Score 36; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      2 LLLPLQILL 9
Db      1 LLLPLQILL 8

RESULT 9
US-09-918-243-36
; Sequence 36, Application US/09918243
; Patent No. US20020142317A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 36

```

; LENGTH: 9
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CHAIN
 ; OTHER INFORMATION: Residues 4-12 of the SCCE protein
 US-09-918-243-36

Query Match 90.0%; Score 36; DB 9; Length 9;
 Best Local Similarity 100.0%; Pred. No. 7.1e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQLL 8
 |||||
 Db 2 LLLPLQLL 9

RESULT 10
 US-09-905-083-35
 ; Sequence 35, Application US/09905083
 ; Patent No. US20020146708A1
 ; GENERAL INFORMATION:
 ; APPLICANT: O'Brien, Timothy J.
 ; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
 ; FILE REFERENCE: D6223CIP/C/Div
 ; CURRENT APPLICATION NUMBER: US/09/905,083
 ; PRIOR FILING DATE: 2001-07-13
 ; PRIOR APPLICATION NUMBER: US 09/502,600
 ; PRIOR FILING DATE: 2000-02-11
 ; NUMBER OF SEQ ID NOS: 136
 ; SEQ ID NO 35
 ; LENGTH: 9
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CHAIN
 ; OTHER INFORMATION: Residues 6-14 of the SCCE protein
 US-09-905-083-35

Query Match 90.0%; Score 36; DB 9; Length 9;
 Best Local Similarity 100.0%; Pred. No. 7.1e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LLLPLQLL 9
 |||||
 Db 1 LLLPLQLL 8

RESULT 11
 US-09-905-083-36
 ; Sequence 36, Application US/09905083
 ; Patent No. US20020146708A1
 ; GENERAL INFORMATION:
 ; APPLICANT: O'Brien, Timothy J.
 ; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
 ; FILE REFERENCE: D6223CIP/C/Div
 ; CURRENT APPLICATION NUMBER: US/09/905,083
 ; PRIOR FILING DATE: 2001-07-13
 ; PRIOR APPLICATION NUMBER: US 09/502,600
 ; PRIOR FILING DATE: 2000-02-11
 ; NUMBER OF SEQ ID NOS: 136
 ; SEQ ID NO 36
 ; LENGTH: 9
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CHAIN
 ; OTHER INFORMATION: Residues 4-12 of the SCCE protein
 US-09-905-083-36

Query Match 90.0%; Score 36; DB 9; Length 9;

Best Local Similarity 100.0%; Pred. No. 7.1e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLPLQLL 8
 |||||
 Db 2 LLLPLQLL 9

RESULT 12
 US-09-864-761-46097
 ; Sequence 46097, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Aemica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 46097
 ; LENGTH: 23
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AJ246003.1
 ; OTHER INFORMATION: EXPRESSED IN RETAL LIVER, SIGNAL = 1.6
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
 ; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
 US-09-864-761-46097

Query Match 85.0%; Score 34; DB 9; Length 23;
Best Local Similarity 87.5%; Pred. No. 20;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LLLPLQILL 9
Db 11 LLLPLQILL 18
|||||:|

RESULT 13
US-09-956-622A-39
; Sequence 39 Application US/09956622A
; Publication No. US20030091973A1
; GENERAL INFORMATION:
; APPLICANT: Horesovsky, Gregory J
; APPLICANT: No. US20030091973A1 II, L. Staton
; APPLICANT: Raha, Debasish
; TITLE OF INVENTION: Method of Identifying Osteoregenerative Agents Using
; TITLE OF INVENTION: Differential Gene Expression
; FILE REFERENCE: 21402-445
; CURRENT APPLICATION NUMBER: US/09/956,622A
; CURRENT FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/233,579
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-956-622A-39

Query Match 85.0%; Score 34; DB 10; Length 201;
Best Local Similarity 77.8%; Pred. No. 1.7e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLLPLQILL 9
Db 6 LLLPLRLLL 14
|||||:|

RESULT 14
US-10-156-761-11235
; Sequence 11235, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11235
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11235

Query Match 82.5%; Score 33; DB 14; Length 516;
Best Local Similarity 77.8%; Pred. No. 6.6e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLLPLQILL 9

Db 373 LLLPLQILL 381
|||||:|

RESULT 15
US-10-096-241-24
; Sequence 24, Application US/10096241
; Publication No. US20020127594A1
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Busfield, Samantha J.
; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES
; AND USES THEREFOR
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/096,241
; FILING DATE: 12-Mar-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/699,591
; FILING DATE: 19-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Faese, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07334/022001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: <unknown>
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-10-096-241-24

Query Match 80.0%; Score 32; DB 13; Length 21;
Best Local Similarity 77.8%; Pred. No. 41;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LLLPLQILL 9
Db 8 LLLPLALL 16
|||||:|

Search completed: March 1, 2004, 18:08:50
Job time : 25.1111 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:21:46 ; Search time 11.8899 Seconds
(without alignments)
39.081 Million cell updates/sec

Title: US-09-905-083-31
Perfect score: 49
Sequence: 1 KXNEYTVHL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgm2_6/ptodata/2/iaa/5A.COMB.pep.*
2: /cgm2_6/ptodata/2/iaa/5B.COMB.pep.*
3: /cgm2_6/ptodata/2/iaa/6A.COMB.pep.*
4: /cgm2_6/ptodata/2/iaa/6B.COMB.pep.*
5: /cgm2_6/ptodata/2/iaa/PCTUS.COMB.pep.*
6: /cgm2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	100.0	9	3	US-09-502-600-31
2	49	100.0	9	3	US-09-502-600-89
3	49	100.0	9	3	US-09-502-600-109
4	49	100.0	9	4	US-09-918-243-31
5	49	100.0	9	4	US-09-918-243-89
6	49	100.0	9	4	US-09-918-243-109
7	49	100.0	144	4	US-09-618-259-4
8	49	100.0	154	3	US-09-261-416-7
9	49	100.0	224	3	US-08-944-483-33
10	49	100.0	225	2	US-08-557-146-12
11	49	100.0	225	2	US-09-027-337-4
12	49	100.0	225	2	US-09-154-344-12
13	49	100.0	225	4	US-09-644-600-4
14	49	100.0	225	4	US-09-654-600A-4
15	49	100.0	253	2	US-08-557-146-2
16	49	100.0	253	2	US-08-824-874-3
17	49	100.0	253	2	US-09-154-344-2
18	49	100.0	253	3	US-09-930-188-2
19	49	100.0	253	3	US-09-210-084-3
20	49	100.0	253	4	US-09-764-762-3
21	49	100.0	253	5	PCT-US98-04294-2
22	39	79.6	9	3	US-09-502-600-122
23	39	79.6	9	4	US-09-918-243-122
24	37	75.5	9	4	US-09-502-600-93
25	37	75.5	9	4	US-09-918-243-93
26	37	75.5	463	4	US-09-540-236-2942
27	33	67.3	204	4	US-09-323-872A-13

28	33	67.3	204	4	US-09-072-433-22	Sequence 22, Appl
29	33	67.3	366	4	US-09-134-001C-5502	Sequence 5502, Ap
30	33	67.3	799	3	US-08-909-954-4	Sequence 4, Appl
31	33	67.3	804	3	US-08-909-954-2	Sequence 2, Appl
32	32	65.3	60	4	US-09-134-001C-5184	Sequence 5184, Ap
33	32	65.3	116	4	US-09-732-210-882	Sequence 882, App
34	32	65.3	121	4	US-09-091-725-49	Sequence 49, Appl
35	32	65.3	792	4	US-09-134-000C-5895	Sequence 5895, Ap
36	31	63.3	258	4	US-09-134-001C-4764	Sequence 4764, Ap
37	31	63.3	575	3	US-08-913-805A-2	Sequence 2, Appl
38	31	63.3	575	3	US-09-442-629-2	Sequence 2, Appl
39	31	63.3	756	3	US-08-982-785A-2	Sequence 2, Appl
40	31	63.3	973	4	US-09-107-532A-4810	Sequence 4810, Ap
41	31	63.3	1084	4	US-09-227-725A-3	Sequence 3, Appl
42	31	63.3	1183	2	US-08-447-031A-2	Sequence 2, Appl
43	30	61.2	28	5	PCT-US94-05150-12	Sequence 12, Appl
44	30	61.2	42	5	PCT-US94-05150-17	Sequence 17, Appl
45	30	61.2	146	4	US-09-489-039A-11204	Sequence 11204, A

ALIGNMENTS

RESULT 1
US-09-502-600-31
; Sequence 31, Application US/09502600A
; Patent No. 6294344
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP-C
; CURRENT FILING DATE: 2000-02-11
; CURRENT APPLICATION NUMBER: US/09/502,600A
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 31
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 72-80 of the SCCE protein
US-09-502-600-31

Query Match 100.0%; Score 49; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KXNEYTVHL 9

Db 1 KXNEYTVHL 9

RESULT 2
US-09-502-600-89
; Sequence 89, Application US/09502600A
; Patent No. 6294344
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP-C
; CURRENT FILING DATE: 2000-02-11
; CURRENT APPLICATION NUMBER: US/09/502,600A
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 89
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

; OTHER INFORMATION: Residues 72-80 of the SCCE protein
US-09-502-600-89

Query Match 100.0%; Score 49; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKNQYTVHL 9
Db 1 KKNQYTVHL 9

RESULT 3
US-09-502-600-109

; Sequence 109, Application US/09502600A
; Patent No. 6294344

; GENERAL INFORMATION:

; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP-C
; CURRENT FILING DATE: 2000-02-11
; CURRENT APPLICATION NUMBER: US/09/502,600A
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 109
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

; OTHER INFORMATION: Residues 72-80 of the SCCE protein
US-09-502-600-109

Query Match 100.0%; Score 49; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKNQYTVHL 9
Db 1 KKNQYTVHL 9

RESULT 4

US-09-918-243-31

; Sequence 31, Application US/09918243

; Patent No. 6627403

; GENERAL INFORMATION:

; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 31
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 72-80 of the SCCE protein
US-09-918-243-31

Query Match 100.0%; Score 49; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKNQYTVHL 9
Db 1 KKNQYTVHL 9

Db 1 KKNQYTVHL 9

RESULT 5

US-09-918-243-89

; Sequence 89, Application US/09918243

; Patent No. 6627403

; GENERAL INFORMATION:

; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 89
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 72-80 of the SCCE protein
US-09-918-243-89

Query Match 100.0%; Score 49; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKNQYTVHL 9
Db 1 KKNQYTVHL 9

RESULT 6

US-09-918-243-109

; Sequence 109, Application US/09918243

; Patent No. 6627403

; GENERAL INFORMATION:

; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 109
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 72-80 of the SCCE protein
US-09-918-243-109

Query Match 100.0%; Score 49; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKNQYTVHL 9
Db 1 KKNQYTVHL 9

RESULT 7

US-09-618-259-4

; Sequence 4, Application US/09618259

; Patent No. 6642013


```

; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: No. 6642013el Extracellular Serine Protease
; FILE REFERENCE: D6020CIP2
; CURRENT APPLICATION NUMBER: US/09/618,259
; CURRENT FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: US 09/127,444
; PRIOR FILING DATE: 1998-08-21
; NUMBER OF SEQ ID NOS: 72
; SEQ ID NO 4
; LENGTH: 144
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; NAME/KEY: DOMAIN
; OTHER INFORMATION: Amino acid sequence of stratum corneum chymotryptic
; OTHER INFORMATION: enzyme (scce) catalytic domain
US-09-618-259-4

Query Match          100.0%; Score 49; DB 4; Length 144;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KMNEYTVHL 9
Db 9 KMNEYTVHL 17

RESULT 8
US-09-261-416-7
; Sequence 7, Application US/09261416A
; Patent No. 6291663
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: TADG-12: A No. 6291663el Transmembrane Serine Protease
; FILE REFERENCE: D6192
; CURRENT APPLICATION NUMBER: US/09/261,416A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 7
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Serine protease catalytic domain of stratum corneum
; OTHER INFORMATION: chymotryptic enzyme (Scce) homologous to similar
; OTHER INFORMATION: domain in TADG-12
US-09-261-416-7

```

```

Query Match          100.0%; Score 49; DB 3; Length 154;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 KMNEYTVHL 9
Db 18 KMNEYTVHL 26

```

```

RESULT 9
US-08-944-483-33
; Sequence 33, Application US/08944493
; Patent No. 6232456
; GENERAL INFORMATION:
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STEWART, KENT D.

```

```

; APPLICANT: STROUPE, STEVEN D.
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,483
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6183.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 224 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6232456e
US-08-944-483-33

Query Match          100.0%; Score 49; DB 3; Length 224;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KMNEYTVHL 9
Db 43 KMNEYTVHL 51

RESULT 10
US-08-557-146-12
; Sequence 12, Application US/08557146
; Patent No. 5834290
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hansson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/557,146
; FILING DATE: 14-DEC-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 225 amino acids
; TYPE: amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
; US-08-557-146-12

Query Match 100.0%; Score 49; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNEYTVHL 9
Db 44 KNEYTVHL 52

RESULT 11
US-09-027-337-4
; Sequence 4, Application US/09027337B
; Patent No. 5972616
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotooshi
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease Overexpressed in
; TITLE OF INVENTION: TADG-15: Breast and Ovarian Carcinomas
; FILE REFERENCE: D6064
; CURRENT APPLICATION NUMBER: US/09/027,337B
; CURRENT FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 13
; SEQ ID NO 4
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Serine protease catalytic domain of Scce homologous to
; OTHER INFORMATION: similar domain in TADG-15
US-09-027-337-4

Query Match 100.0%; Score 49; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNEYTVHL 9
Db 44 KNEYTVHL 52

RESULT 12
US-09-154-344-12
; Sequence 12, Application US/09154344
; Patent No. 5981256
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hansson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York

```

```

; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/154,344
; FILING DATE: 16-SEP-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/557,146
; FILING DATE: 14-DEC-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 225 amino acids
; TYPE: amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
; US-09-154-344-12

Query Match 100.0%; Score 49; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNEYTVHL 9
Db 44 KNEYTVHL 52

RESULT 13
US-09-644-600-4
; Sequence 4, Application US/09644600
; Patent No. 6451500
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotooshi
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease
; TITLE OF INVENTION: Overexpressed in Carcinomas
; FILE REFERENCE: D6064CIP/D
; CURRENT APPLICATION NUMBER: US/09/644,600
; CURRENT FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/421,213
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: 09/027,337
; PRIOR FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 4
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: SCCE
; US-09-644-600-4

Query Match 100.0%; Score 49; DB 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNEYTVHL 9
Db 44 KNEYTVHL 52

```

RESULT 14
US-09-654-600A-4
; Sequence 4, Application US/09654600A
; Patent No. 6649741
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotochi
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease
; TITLE OF INVENTION: Overexpressed in Carcinomas
; FILE REFERENCE: D6064CIP/D
; CURRENT APPLICATION NUMBER: US/09/654,600A
; CURRENT FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/421,213
; 09/027,337
; PRIOR FILING DATE: 1999-10-20
; 1998-02-20
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 4
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: SCCE
US-09-654-600A-4
Query Match 100.0%; Score 49; DB 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KXNEYTVHL 9
DB 44 KXNEYTVHL 52
RESULT 15
US-08-557-146-2
; Sequence 2, Application US/08557146
; Patent No. 5834290
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hansson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,146
; FILING DATE: 14-DEC-1995
; CLASSIFICATION: 42A
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid

TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-557-146-2
Query Match 100.0%; Score 49; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KXNEYTVHL 9
DB 72 KXNEYTVHL 80
Search completed: March 1, 2004, 17:38:22
Job time : 12.8889 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:16:55 ; Search time 45.5556 Seconds
(without alignments)
55.820 Million cell updates/sec

Title: US-09-905-083-31

Perfect score: 49

Sequence: 1 KNEYTVHL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	100.0	9	AAE08294	Human str
2	49	100.0	9	AAE08236	Human str
3	49	100.0	97	ADA05740	Human NOV
4	49	100.0	136	ABG23378	Novel hum
5	49	100.0	181	ADA05738	Human NOV
6	49	100.0	138	ADA05736	Human NOV
7	49	100.0	224	ADA05744	Human NOV
8	49	100.0	225	AAE08236	Human str
9	49	100.0	247	ADA05742	Human NOV
10	49	100.0	250	ADA05732	Human NOV
11	49	100.0	252	ADA05734	Human NOV
12	49	100.0	253	AAE08236	Human str
13	49	100.0	253	AAE08236	Human str
14	49	100.0	253	AAE08236	Human str
15	49	100.0	253	AAE08236	Human str
16	49	100.0	253	AAE08236	Human str
17	49	100.0	253	AAE08236	Human str
18	49	100.0	253	AAE08236	Human str
19	49	100.0	253	AAE08236	Human str
20	49	100.0	253	AAE08236	Human str
21	49	100.0	253	AAE08236	Human str
22	49	100.0	253	AAE08236	Human str
23	39	79.6	9	AAE08298	Human str
24	37	75.5	9	AAE08298	Human str
25	37	75.5	243	ABB84419	Bovine SC

26	37	75.5	334	4	AAU37656	Streptococ
27	37	75.5	334	6	ABU46099	Protein e
28	37	75.5	340	6	ABU35220	Protein e
29	36	73.5	249	5	ABB84420	Porcine S
30	36	73.5	753	5	ABP25540	Streptococ
31	35	71.4	159	5	ABB48701	Listeria
32	35	71.4	434	4	AAE08085	Human pro
33	35	71.4	685	4	AAE08085	Human pro
34	35	71.4	685	4	ABG18060	Novel hum
35	35	71.4	744	4	ABG18061	Novel hum
36	34	69.4	128	5	ABP08668	Human ORF
37	34	69.4	270	6	ABR64244	Angiogene
38	34	69.4	270	6	ABU03733	Human exp
39	34	69.4	270	6	ABU03734	Human exp
40	34	69.4	270	6	ABU03735	Human exp
41	34	69.4	270	6	ABU03737	Human exp
42	34	69.4	270	6	ABU03738	Human exp
43	34	69.4	270	6	AAO30066	Human rep
44	34	69.4	421	5	ABB48147	Listeria
45	34	69.4	421	6	ABU32976	Protein e

ALIGNMENTS

RESULT 1

AAE08294
ID AAE08294 standard; peptide; 9 AA.
XX
AC AAE08294;
XX
DT 01-NOV-2001 (first entry)
XX
DE Human stratum corneum chymotrypsin enzyme peptide #59 (residues 72-80).
XX
KW Stratum corneum chymotrypsin enzyme; SCCE; cytostatic; vaccine; tumour;
KW cancer; ovarian; breast; lung; colon; prostate; carcinoma; human;
KW antisense therapy; malignant hyperplasia.
XX
OS Homo sapiens.
XX
PN WO200159158-A1.
XX
PD 16-AUG-2001.
XX
PF 07-FEB-2001; 2001WO-US003977.
XX
PR 11-FEB-2000; 2000US-00502600.
XX
(UYAR-) UNIV ARKANSAS.
XX
O'brien TU;
XX
WPI; 2001-514676/56.
XX
Diagnosing cancer comprises detecting stratum corneum chymotrypsin enzyme.
XX
Disclosure; Page 115; 127pp; English.

The invention relates to diagnosing cancer especially ovarian cancer, by screening for stratum corneum chymotrypsin enzyme (SCCE). Proteases are considered to be an integral part of tumour growth and metastasis, and therefore markers indicative of their presence or absence are useful for the diagnosis of cancer. The method is useful for diagnosing cancer, carcinoma and malignant hyperplasia. The SCCE oligonucleotide may be used to treat a cancer selected from ovarian, breast, lung, colon, prostate and other cancers in which SCCE is overexpressed. The present sequence is human SCCE peptide

XX Sequence 9 AA;

Query Match 100.0%; Score 49; DB 4; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KQNEYTVHL 9

Db 1 KQNEYTVHL 9

RESULT 2

AAE08236

ID AAE08236 standard; peptide; 9 AA.

XX AC AAE08236;

XX DF 01-NOV-2001 (first entry)

XX DE Human stratum corneum chymotrypsin enzyme peptide #1 (residues 72-80).

XX KW Stratum corneum chymotrypsin enzyme; SCCE; cytostatic; vaccine; tumour;
KW cancer; ovarian; breast; lung; colon; prostate; carcinoma; human;
KW anisense therapy; malignant hyperplasia.

XX OS Homo sapiens.

XX PN WO200159158-A1.

XX PD 16-AUG-2001.

XX PF 07-FEB-2001; 2001WO-US003977.

XX PR 11-FEB-2000; 2000US-00502600.

XX FA (UYAR-) UNIV ARKANSAS.

XX PI O'Brien TJ;

XX DR WPI; 2001-514676/56.

XX PT Diagnosing cancer comprises detecting stratum corneum chymotrypsin enzyme.

XX PS Claim 25; Page 102; 127pp; English.

XX CC The invention relates to diagnosing cancer especially ovarian cancer, by screening for stratum corneum chymotrypsin enzyme (SCCE). Proteases are considered to be an integral part of tumour growth and metastasis, and therefore, markers indicative of their presence or absence are useful for the diagnosis of cancer. The method is useful for diagnosing cancer, carcinoma and malignant hyperplasia. The SCCE oligonucleotide may be used to treat a cancer selected from ovarian, breast, lung, colon, prostate and other cancers in which SCCE is overexpressed. The present sequence is human SCCE peptide

XX SQ Sequence 9 AA;

Query Match

Best Local Similarity 100.0%; Pred. No. 1.4e+06;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KQNEYTVHL 9

Db 1 KQNEYTVHL 9

RESULT 3

ADA05740

ID ADA05740 standard; protein; 97 AA.

XX AC ADA05740;

XX DT 06-NOV-2003 (first entry)

XX DE Human NOV18e protein SEQ ID NO:100.

XX KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
KW immunomodulator; cytostatic; nootropic; neuroprotective;
KW antiparkinsonian; antilipemic; gene therapy; human disease;
KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW immune disorder; haematopoietic disorder; dyslipidaemia;
XX OS Homo sapiens.
XX PN WO2003029424-A2.
XX PD 10-APR-2003.
XX PF 02-OCT-2002; 2002WO-US031373.
XX PR 02-OCT-2001; 2001US-0326483P.
XX PR 05-OCT-2001; 2001US-0327435P.
XX PR 05-OCT-2001; 2001US-0327449P.
XX PR 09-OCT-2001; 2001US-0327917P.
XX PR 09-OCT-2001; 2001US-0328029P.
XX PR 09-OCT-2001; 2001US-0328044P.
XX PR 09-OCT-2001; 2001US-0328056P.
XX PR 12-OCT-2001; 2001US-0328849P.
XX PR 15-OCT-2001; 2001US-0329414P.
XX PR 17-OCT-2001; 2001US-0330142P.
XX PR 18-OCT-2001; 2001US-0330309P.
XX PR 22-OCT-2001; 2001US-0341058P.
XX PR 24-OCT-2001; 2001US-0339266P.
XX PR 24-OCT-2001; 2001US-0343629P.
XX PR 29-OCT-2001; 2001US-0349575P.
XX PR 01-NOV-2001; 2001US-0346357P.
XX PR 17-APR-2002; 2002US-0373260P.
XX PR 19-APR-2002; 2002US-0373815P.
XX PR 19-APR-2002; 2002US-0373817P.
XX PR 19-APR-2002; 2002US-0373826P.
XX PR 19-APR-2002; 2002US-0373884P.
XX PR 22-APR-2002; 2002US-0374977P.
XX PR 16-MAY-2002; 2002US-0381037P.
XX PR 16-MAY-2002; 2002US-0381038P.
XX PR 17-MAY-2002; 2002US-0381042P.
XX PR 28-MAY-2002; 2002US-0381642P.
XX PR 29-MAY-2002; 2002US-0383656P.
XX PR 25-JUN-2002; 2002US-0383831P.
XX PR 01-OCT-2002; 2002US-0391335P.
XX PR 01-OCT-2002; 2002US-00262511.
XX (CURA-) CURAGEN CORP.
XX PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
PI Ort T, Gorman L, Zernusen BD, Anderson DW, Zhong M, Catterton E;
PI Ji W, Miller CE, Rastelli L, Stone DU, Pena CBA, Shenoy SG;
PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;
PI Eisen AJ, Gargolli EA, Rieger DK, Spaderna SK;
XX WPI: 2003-381626/36.
XX N-PSDB; ADA05739.
XX PT New NOVX polypeptides and nucleic acids, useful for diagnosing,
XX preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
XX cancer or dyslipidemia, and in chromosome mapping, tissue typing or
XX pharmacogenomics.
XX PS Claim 1; Page 171; 586pp; English.
XX CC The present invention describes NOVX proteins, where X can be 1 to 55
XX (e.g. NOV1). Also described: (1) a composition comprising a polypeptide
XX described above and a carrier; (2) a kit comprising, in one or more
XX containers, the composition described above; (3) an isolated nucleic acid
XX molecule which encodes a NOVX protein of the invention; (4) a vector
XX comprising the nucleic acid molecule described above; (5) a cell
XX comprising the above vector; (6) an antibody that immunospecifically

CC binds to the polypeptide described above; (7) methods for determining the
 CC presence or amount of the above polypeptide or nucleic acid molecule in a
 CC sample; (8) methods for determining the presence of or predisposition to
 CC a disease associated with altered levels of expression of the above
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
 CC method of identifying an agent that binds to the polypeptide described
 CC above; (10) a method for identifying a potential therapeutic agent for
 CC use in treating a pathology that is related to an aberrant expression or
 CC aberrant physiological interactions of the polypeptide; (11) a method of
 CC screening for a modulator of activity or of latency or predisposition to
 CC a pathology associated with the polypeptide; (12) a method for modulating
 CC the activity of the polypeptide described above; (13) methods of treating
 CC or preventing a pathology associated with the above polypeptide in a
 CC mammal; and (14) a method for producing the above polypeptide in a
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,
 CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian
 CC and antilipemic activities, and can be used in gene therapy. The
 CC polypeptide is useful in manufacturing a medicament for treating a
 CC syndrome associated with a human disease. The polypeptide or the nucleic
 CC acid molecule may be used to diagnose, treat or prevent metabolic
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
 CC disease, immune disorders, haematopoietic disorders and various
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation
 CC probes, in chromosome mapping, tissue typing, preventive medicine and
 CC pharmacogenomics. The present sequence represents a human NOVX from the
 CC present invention.

SQ Sequence 97 AA;

Query Match 100.0%; Score 49; DB 6; Length 97;
 Best Local Similarity 100.0%; Pred. No. 0.081;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KMEYTVHL 9
 |||||
 Db 53 KMEYTVHL 61

RESULT 4

ID ABG23378
 ID ABG23378 standard; protein; 136 AA.

XX
 AC ABG23378;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #23369.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Dmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS87565.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

XX Claim 20; SEQ ID NO 53737; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 136 AA;

Query Match 100.0%; Score 49; DB 4; Length 136;

Best Local Similarity 100.0%; Pred. No. 0.12; 0; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0;

QY 1 KMEYTVHL 9
 |||||
 Db 78 KMEYTVHL 86

RESULT 5

ID ADA05738
 ID ADA05738 standard; protein; 181 AA.

XX
 AC ADA05738;

XX 06-NOV-2003 (first entry)

XX Human NOV18d protein SEQ ID NO:98.

XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
 KW immunomodulator; cytostatic; nootropic; neuroprotective;
 KW antiparkinsonian; antilipemic; gene therapy; human disease;
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
 KW neurodegenerative disorder; Alzheimer's disease; parkinson's disease;
 KW immune disorder; haematopoietic disorder; dyslipidaemia.

XX Homo sapiens.

XX WO2003029424-A2.

XX 10-APR-2003.

XX 02-OCT-2002; 2002WO-US031373.

PR 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327435P.

PR 05-OCT-2001; 2001US-0327449P.

PR 09-OCT-2001; 2001US-0327917P.

PR 09-OCT-2001; 2001US-0328029P.

PR 09-OCT-2001; 2001US-0328044P.

PR 09-OCT-2001; 2001US-0328056P.

PR 12-OCT-2001; 2001US-0328849P.

PR 15-OCT-2001; 2001US-0329414P.

PR 17-OCT-2001; 2001US-0330142P.

PR 18-OCT-2001; 2001US-0330309P.

PR 22-OCT-2001; 2001US-0341058P.
 PR 24-OCT-2001; 2001US-0339266P.
 PR 24-OCT-2001; 2001US-0343629P.
 PR 29-OCT-2001; 2001US-0349575P.
 PR 01-NOV-2001; 2001US-0346357P.
 PR 17-APR-2002; 2002US-0373260P.
 PR 19-APR-2002; 2002US-0373815P.
 PR 19-APR-2002; 2002US-0373817P.
 PR 19-APR-2002; 2002US-0373826P.
 PR 19-APR-2002; 2002US-0373884P.
 PR 22-APR-2002; 2002US-0374977P.
 PR 16-MAY-2002; 2002US-0381037P.
 PR 16-MAY-2002; 2002US-0381038P.
 PR 16-MAY-2002; 2002US-0381042P.
 PR 17-MAY-2002; 2002US-0381642P.
 PR 28-MAY-2002; 2002US-0383656P.
 PR 29-MAY-2002; 2002US-0383831P.
 PR 25-JUN-2002; 2002US-0391335P.
 PR 01-OCT-2002; 2002US-00262511.
 XX

PA (CURA-) CURAGEN CORP.

XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
 PI Patturajan M, Spytsek KA, Edinger SR, Ellerman K, Malyankar UM;
 PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;
 PI Ji W, Miller CE, Rasstelli L, Stone DU, Pena CEA, Shenoy SG;
 PI Shimkets RA, Rothenberg MB, Leach MD, Agee ML, Berghs C, Dipippo VA;
 PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;
 XX

DR WPI; 2003-381626/36.
 DR N-PSDB; ADA05737.
 XX

PT New NOVX polypeptides and nucleic acids, useful for diagnosing,
 PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
 PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.
 XX

PS Claim 1; Page 171; 586pp; English.

XX The present invention describes NOVX proteins, where X can be 1 to 55
 CC (e.g. NOVX). Also described: (1) a composition comprising a polypeptide
 CC described above and a carrier; (2) a kit comprising, in one or more
 CC containers, the composition described above; (3) an isolated nucleic acid
 CC molecule which encodes a NOVX protein of the invention; (4) a vector
 CC comprising the nucleic acid molecule described above; (5) a cell
 CC comprising the above vector; (6) an antibody that immunospecifically
 CC binds to the polypeptide described above; (7) methods for determining the
 CC presence or amount of the above polypeptide or nucleic acid molecule in a
 CC sample; (8) methods for determining the presence of or predisposition to
 CC a disease associated with altered levels of expression of the above
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
 CC method of identifying an agent that binds to the polypeptide described
 CC above; (10) a method for identifying a potential therapeutic agent for
 CC use in treating a pathology that is related to an aberrant expression or
 CC aberrant physiological interactions of the polypeptide; (11) a method of
 CC screening for a modulator of activity or of latency or predisposition to
 CC a pathology associated with the polypeptide; (12) a method for modulating
 CC the activity of the polypeptide described above; (13) methods of treating
 CC or preventing a pathology associated with the above polypeptide in a
 CC mammal; and (14) a method for producing the above polypeptide. NOVX
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,
 CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian
 CC and antilipidemic activities, and can be used in gene therapy. The
 CC polypeptide is useful in manufacturing a medicament for treating a
 CC syndrome associated with a human disease. The polypeptide or the nucleic
 CC acid molecule may be used to diagnose, treat or prevent metabolic
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
 CC disease, immune disorders, haematopoietic disorders and various
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation
 CC probes, in chromosome mapping, tissue typing, preventive medicine and
 CC pharmacogenomics. The present sequence represents a human NOVX from the
 CC present invention.

XX SQ Sequence 181 AA;
 Query Match 100.0%; Score 49; DB 6; Length 181;
 Best Local Similarity 100.0%; Pred. No. 0.16; Mismatches 0; Gaps 0;
 Matches 9; Conservative 0; Indels 0;
 QY 1 KXNEYTVHL 9
 Db 53 KXNEYTVHL 61

RESULT 6

ADA05736
 ID ADA05736 standard; protein; 198 AA.

XX ADA05736;

XX 06-NOV-2003 (first entry)

XX Human NOV18c protein SEQ ID NO:96.

XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
 KW immunomodulator; cytostatic; nootropic; neuroprotective;
 KW antiparkinsonian; antilipidemic; gene therapy; human disease;
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; dyslipidaemia.

XX Homo sapiens.

XX WO2003029424-A2.

XX 10-APR-2003.

XX 02-OCT-2002; 2002WO-US031373.

XX 02-OCT-2001; 2001US-0326483P.

XX 05-OCT-2001; 2001US-0327435P.

XX 09-OCT-2001; 2001US-0327917P.

XX 09-OCT-2001; 2001US-0328029P.

XX 09-OCT-2001; 2001US-0328044P.

XX 09-OCT-2001; 2001US-0328056P.

XX 12-OCT-2001; 2001US-0328649P.

XX 15-OCT-2001; 2001US-0329414P.

XX 17-OCT-2001; 2001US-0330142P.

XX 18-OCT-2001; 2001US-0330309P.

XX 22-OCT-2001; 2001US-0341058P.

XX 24-OCT-2001; 2001US-0339266P.

XX 29-OCT-2001; 2001US-0349575P.

XX 01-NOV-2001; 2001US-0346357P.

XX 17-APR-2002; 2002US-0373260P.

XX 19-APR-2002; 2002US-0373815P.

XX 19-APR-2002; 2002US-0373817P.

XX 19-APR-2002; 2002US-0373826P.

XX 22-APR-2002; 2002US-0374977P.

XX 16-MAY-2002; 2002US-0381037P.

XX 16-MAY-2002; 2002US-0381038P.

XX 17-MAY-2002; 2002US-0381642P.

XX 28-MAY-2002; 2002US-0383656P.

XX 29-MAY-2002; 2002US-0383831P.

XX 25-JUN-2002; 2002US-0391335P.

XX 01-OCT-2002; 2002US-00262511.

(CURA-) CURAGEN CORP.

XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
 PI Patturajan M, Spytsek KA, Edinger SR, Ellerman K, Malyankar UM;
 PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;

PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;
 PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;
 PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;
 DR WPI; 2003-381626/36.
 DR N-PSDB; ADA05735.
 XX
 XX
 PT New NOVX polypeptides and nucleic acids, useful for diagnosing,
 PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
 PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.
 XX
 PS Claim 1; Page 170; 586pp; English.
 XX
 CC The present invention describes NOVX proteins, where X can be 1 to 55
 CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide
 CC described above and a carrier; (2) a kit comprising, in one or more
 CC containers, the composition described above; (3) an isolated nucleic acid
 CC molecule which encodes a NOVX protein of the invention; (4) a vector
 CC comprising the nucleic acid molecule described above; (5) a cell
 CC comprising the above vector; (6) an antibody that immunospecifically
 CC binds to the polypeptide described above; (7) methods for determining the
 CC presence or amount of the above polypeptide or nucleic acid molecule in a
 CC sample; (8) methods for determining the presence of or predisposition to
 CC a disease associated with altered levels of expression of the above
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
 CC method of identifying an agent that binds to the polypeptide described
 CC above; (10) a method for identifying a potential therapeutic agent for
 CC use in treating a pathology that is related to an aberrant expression or
 CC aberrant physiological interactions of the polypeptide; (11) a method of
 CC screening for a modulator of activity of or latency or predisposition to
 CC a pathology associated with the polypeptide; (12) a method for modulating
 CC the activity of the polypeptide described above; (13) methods of treating
 CC or preventing a pathology associated with the above polypeptide in a
 CC mammal; and (14) a method for producing the above polypeptide. NOVX
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,
 CC immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian
 CC and antilipidemic activities, and can be used in gene therapy. The
 CC polypeptide is useful in manufacturing a medicament for treating a
 CC syndrome associated with a human disease. The polypeptide or the nucleic
 CC acid molecule may be used to diagnose, treat or prevent metabolic
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
 CC disease, immune disorders, haematopoietic disorders and various
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation
 CC probes, in chromosome mapping, tissue typing, preventive medicine and
 CC pharmacogenomics. The present sequence represents a human NOVX from the
 CC present invention.
 XX
 SQ Sequence 198 AA;
 Query Match 100.0%; Score 49; DB 6; Length 198;
 Best Local Similarity 100.0%; Pred. No. 0.18;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KXNEYTVHL 9
 Db 72 KXNEYTVHL 80
 |||||
 |||||
 RESULT 7
 ADA05744
 ID ADA05744 standard; protein; 224 AA.
 XX
 AC ADA05744;
 XX
 DT 06-NOV-2003 (first entry)
 XX
 DE Human NOV18g protein SEQ ID NO:104.
 XX
 XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
 KW immunomodulator; cytostatic; neurotropic; neuroprotective;
 KW antiparkinsonian; antilipidemic; gene therapy; human disease;
 KW

KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; dyslipidaemia.
 XX
 OS Homo sapiens.
 XX
 XX WO2003029424-A2.
 XX
 PD 10-APR-2003.
 XX
 PF 02-OCT-2002; 2002WO-US031373.
 XX
 XX 02-OCT-2001; 2001US-0326483P.
 PR 05-OCT-2001; 2001US-0327435P.
 PR 09-OCT-2001; 2001US-0327449P.
 PR 09-OCT-2001; 2001US-0327917P.
 PR 09-OCT-2001; 2001US-0328029P.
 PR 09-OCT-2001; 2001US-0328044P.
 PR 09-OCT-2001; 2001US-0328056P.
 PR 12-OCT-2001; 2001US-0328849P.
 PR 15-OCT-2001; 2001US-0329414P.
 PR 17-OCT-2001; 2001US-0330142P.
 PR 18-OCT-2001; 2001US-0330309P.
 PR 22-OCT-2001; 2001US-0341058P.
 PR 24-OCT-2001; 2001US-0339256P.
 PR 24-OCT-2001; 2001US-0343629P.
 PR 29-OCT-2001; 2001US-0349576P.
 PR 01-NOV-2001; 2001US-0346357P.
 PR 17-APR-2002; 2002US-0373260P.
 PR 19-APR-2002; 2002US-0373815P.
 PR 19-APR-2002; 2002US-0373817P.
 PR 19-APR-2002; 2002US-0373826P.
 PR 19-APR-2002; 2002US-0373884P.
 PR 22-APR-2002; 2002US-0374977P.
 PR 16-MAY-2002; 2002US-0381037P.
 PR 16-MAY-2002; 2002US-0381038P.
 PR 16-MAY-2002; 2002US-0381042P.
 PR 17-MAY-2002; 2002US-0381642P.
 PR 28-MAY-2002; 2002US-0383656P.
 PR 29-MAY-2002; 2002US-0383831P.
 PR 25-JUN-2002; 2002US-0391335P.
 PR 01-OCT-2002; 2002US-00262511.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
 PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
 PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Cattarotto E;
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;
 PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;
 PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;
 XX
 XX WPI; 2003-381626/36.
 DR N-PSDB; ADA05743.
 XX
 DR New NOVX polypeptides and nucleic acids, useful for diagnosing,
 PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
 PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.
 XX
 PT
 PS Claim 1; Page 172; 586pp; English.
 XX
 CC The present invention describes NOVX proteins, where X can be 1 to 55
 CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide
 CC described above and a carrier; (2) a kit comprising, in one or more
 CC containers, the composition described above; (3) an isolated nucleic acid
 CC molecule which encodes a NOVX protein of the invention; (4) a vector
 CC comprising the nucleic acid molecule described above; (5) a cell
 CC comprising the above vector; (6) an antibody that immunospecifically
 CC binds to the polypeptide described above; (7) methods for determining the
 CC presence or amount of the above polypeptide or nucleic acid molecule in a
 CC sample; (8) methods for determining the presence of or predisposition to
 CC a disease associated with altered levels of expression of the above
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
 CC method of identifying an agent that binds to the polypeptide described
 CC above; (10) a method for identifying a potential therapeutic agent for
 CC use in treating a pathology that is related to an aberrant expression or
 CC aberrant physiological interactions of the polypeptide; (11) a method of
 CC screening for a modulator of activity of or latency or predisposition to
 CC a pathology associated with the polypeptide; (12) a method for modulating
 CC the activity of the polypeptide described above; (13) methods of treating
 CC or preventing a pathology associated with the above polypeptide in a
 CC mammal; and (14) a method for producing the above polypeptide. NOVX
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,
 CC immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian
 CC and antilipidemic activities, and can be used in gene therapy. The
 CC polypeptide is useful in manufacturing a medicament for treating a
 CC syndrome associated with a human disease. The polypeptide or the nucleic
 CC acid molecule may be used to diagnose, treat or prevent metabolic
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
 CC disease, immune disorders, haematopoietic disorders and various
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation
 CC probes, in chromosome mapping, tissue typing, preventive medicine and
 CC pharmacogenomics. The present sequence represents a human NOVX from the
 CC present invention.

CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
 CC method of identifying an agent that binds to the polypeptide described
 CC above; (10) a method for identifying a potential therapeutic agent for
 CC use in treating a pathology that is related to an aberrant expression or
 CC aberrant physiological interactions of the polypeptide; (11) a method of
 CC screening for a modulator of activity or of latency or predisposition to
 CC a pathology associated with the polypeptide; (12) a method for modulating
 CC the activity of the polypeptide described above; (13) methods of treating
 CC or preventing a pathology associated with the above polypeptide in a
 CC mammal; and (14) a method for producing the above polypeptide. NOVX
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,
 CC immunomodulator, cytosstatic, neurotropic, neuroprotective, antiparkinsonian
 CC and antilipaeamic activities, and can be used in gene therapy. The
 CC polypeptide is useful in manufacturing a medicament for treating a
 CC syndrome associated with a human disease. The polypeptide or the nucleic
 CC acid molecule may be used to diagnose, treat or prevent metabolic
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
 CC disease, immune disorders, haematopoietic disorders and various
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation
 CC probes, in chromosome mapping, tissue typing, preventive medicine and
 CC present invention. The present sequence represents a human NOVX from the
 CC
 XX Sequence 224 AA;
 SQ

Query Match 100.0%; Score 49; DB 6; Length 224;
 Best Local Similarity 100.0%; Pred. No. 0.21;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNEVTVHL 9
 DB 58 KNEVTVHL 66
 |||||

RESULT 8
 AAB98502
 ID AAB98502 standard; protein; 225 AA.

XX AC AAB98502;
 XX DT 03-AUG-2001 (first entry)
 XX DE Human Stratum Corneum Chymotryptic Enzyme, SCCE, catalytic domain.
 XX KW Human; TADG-15; cytostatic; vaccine; ovarian tumour; cancer;
 KW tumour antigen-derived gene 15; serine protease;
 KW Stratum Corneum Chymotryptic Enzyme; SCCE.

XX OS Homo sapiens.
 XX PN WO200129056-A1.
 XX PD 26-APR-2001.

XX PF 20-OCT-2000; 2000WO-US029095.

XX PR 20-OCT-1999; 99US-00421213.

XX PA (UYAR-) UNIV ARKANSAS.

XX PI O'brien TJ, Tanimoto H;

XX DR WPI; 2001-381031/40.

XX PT Novel extracellular serine protease, termed tumor antigen-derived gene 15
 PT protein overexpressed in carcinomas and DNA encoding it, for diagnosis,
 PT treatment, prevention of cancer, particularly breast, ovarian cancer.

XX PS Example 10; Fig 1; 130pp; English.

XX CC The present invention relates to human tumour antigen-derived gene 15
 CC (TADG-15) protein and coding sequence (see AAB98502 and AAB98500). TADG-

CC 15 is an extracellular serine protease. It was found that TADG-15 is over
 CC -expressed in ovarian tumours. TADG-15 protein or its fragments of 9-20
 CC residues that lack TADG-15 protease activity are useful for vaccinating
 CC an individual against TADG-15, having, suspected of having or at risk of
 CC getting cancer. Furthermore, the TADG-15 gene can be used as a diagnostic
 CC or therapeutic target in cancer. The present sequence was used in a
 CC sequence homology alignment with the catalytic domain of TADG-15
 XX
 SQ Sequence 225 AA;

Query Match 100.0%; Score 49; DB 4; Length 225;
 Best Local Similarity 100.0%; Pred. No. 0.21;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNEVTVHL 9
 DB 44 KNEVTVHL 52
 |||||

RESULT 9
 ADA05742
 ID ADA05742 standard; protein; 247 AA.

XX AC ADA05742;

XX DT 06-NOV-2003 (first entry)

XX DE Human NOV18f protein SEQ ID NO:102.

XX KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
 KW immunomodulator; cytostatic; neurotropic; neuroprotective;
 KW antiparkinsonian; antilipaeamic; gene therapy; human disease;
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; dyslipidaemia.

XX OS Homo sapiens.

XX PN WO2003029424-A2.

XX PD 10-APR-2003.

XX PF 02-OCT-2002; 2002WO-US031373.

XX PR 02-OCT-2001; 2001US-0326483P.

XX PR 03-OCT-2001; 2001US-0327435P.

XX PR 05-OCT-2001; 2001US-0327449P.

XX PR 09-OCT-2001; 2001US-0327917P.

XX PR 09-OCT-2001; 2001US-0328029P.

XX PR 09-OCT-2001; 2001US-0328044P.

XX PR 09-OCT-2001; 2001US-0328056P.

XX PR 12-OCT-2001; 2001US-0328849P.

XX PR 15-OCT-2001; 2001US-0329414P.

XX PR 17-OCT-2001; 2001US-0330142P.

XX PR 18-OCT-2001; 2001US-0330309P.

XX PR 22-OCT-2001; 2001US-0341058P.

XX PR 24-OCT-2001; 2001US-0339266P.

XX PR 24-OCT-2001; 2001US-0343629P.

XX PR 29-OCT-2001; 2001US-0349575P.

XX PR 01-NOV-2001; 2001US-0346357P.

XX PR 17-APR-2002; 2002US-0373260P.

XX PR 19-APR-2002; 2002US-0373815P.

XX PR 19-APR-2002; 2002US-0373817P.

XX PR 19-APR-2002; 2002US-0373826P.

XX PR 19-APR-2002; 2002US-0373844P.

XX PR 22-APR-2002; 2002US-0374977P.

XX PR 16-MAY-2002; 2002US-0381037P.

XX PR 16-MAY-2002; 2002US-0381038P.

XX PR 16-MAY-2002; 2002US-0381042P.

XX PR 17-MAY-2002; 2002US-0381642P.

XX PR 28-MAY-2002; 2002US-0383656P.

XX PR 29-MAY-2002; 2002US-0383831P.

XX PR 25-JUN-2002; 2002US-0391335P.

01-OCT-2002; 2002US-00262511.
(CURA-) CURAGEN CORP.

Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X; Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM; Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E; Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG; Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Bergins C, Dipippo VA; Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK; WPI; 2003-381626/36.
N-ESDB; ADA05741.

New NOVX polypeptides and nucleic acids, useful for diagnosing, preventing or treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or dyslipidemia, and in chromosome mapping, tissue typing or pharmacogenomics.

Claim 1; Page 172; 586pp; English.

The present invention describes NOVX proteins, where X can be 1 to 55 (e.g. NOV1). Also described: (1) a composition comprising a polypeptide described above and a carrier; (2) a kit comprising, in one or more containers, the composition described above; (3) an isolated nucleic acid molecule which encodes a NOVX protein of the invention; (4) a vector comprising the nucleic acid molecule described above; (5) a cell comprising the above vector; (6) an antibody that immunospecifically binds to the polypeptide described above; (7) methods for determining the presence or amount of the above polypeptide or nucleic acid molecule in a sample; (8) methods for determining the presence of or predisposition to a disease associated with altered levels of expression of the above polypeptide or nucleic acid molecule in a first mammalian subject; (9) a method of identifying an agent that binds to the polypeptide described above; (10) a method for identifying a potential therapeutic agent for use in treating a pathology that is related to an aberrant expression or aberrant physiological interactions of the polypeptide; (11) a method of screening for a modulator of activity or of latency or predisposition to a pathology associated with the polypeptide; (12) a method for modulating the activity of the polypeptide described above; (13) methods of treating or preventing a pathology associated with the above polypeptide in a mammal; and (14) a method for producing the above polypeptide in a sequences have antidiabetic, anorectic, antibacterial, virucide, immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian and antilipidemic activities, and can be used in gene therapy. The polypeptide is useful in manufacturing a medicament for treating a acid molecule may be used to diagnose, treat or prevent metabolic disorders such as diabetes or obesity, infections, cachexia, cancer, neurodegenerative disorders such as Alzheimer's disease or Parkinson's disease, immune disorders, haematopoietic disorders and various dyslipidaemias. The nucleic acids can also be used as hybridisation probes, in chromosome mapping, tissue typing, preventive medicine and pharmacogenomics. The present sequence represents a human NOVX from the present invention.

Sequence 247 AA;
Query Match 100.0%; Score 49; DB 6; Length 247;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KXNEYTVHL 9
Db 58 KXNEYTVHL 66
|||||

RESULT 10
ADA05732
ID ADA05732 standard; protein; 250 AA.
XX
AC ADA05732;
XX

06-NOV-2003 (first entry)
Human NOV18a protein SEQ ID NO:92.

human; NOVX; antidiabetic; anorectic; antibacterial; virucide; immunomodulator; cytostatic; neurotropic; neuroprotective; antiparkinsonian; antilipidemic; gene therapy; human disease; metabolic disorder; diabetes; obesity; infection; cachexia; cancer; neurodegenerative disorder; Alzheimer's disease; Parkinson's disease; immune disorder; haematopoietic disorder; dyslipidaemia.

Homo sapiens.
OS
XX
WC2003029424-A2.
XX
10-APR-2003.
XX
02-OCT-2002; 2002WO-US0311373.
XX
02-OCT-2001; 2001US-0326483P.
PR
05-OCT-2001; 2001US-0327435P.
PR
09-OCT-2001; 2001US-0327449P.
PR
09-OCT-2001; 2001US-0327917P.
PR
09-OCT-2001; 2001US-0328029P.
PR
09-OCT-2001; 2001US-0328044P.
PR
09-OCT-2001; 2001US-0328056P.
PR
12-OCT-2001; 2001US-0328849P.
PR
15-OCT-2001; 2001US-0329414P.
PR
17-OCT-2001; 2001US-0330142P.
PR
18-OCT-2001; 2001US-0330309P.
PR
22-OCT-2001; 2001US-0341058P.
PR
24-OCT-2001; 2001US-0339266P.
PR
24-OCT-2001; 2001US-0343629P.
PR
29-OCT-2001; 2001US-0349575P.
PR
01-NOV-2001; 2001US-0346357P.
PR
17-APR-2002; 2002US-0373260P.
PR
19-APR-2002; 2002US-0373815P.
PR
19-APR-2002; 2002US-0373817P.
PR
19-APR-2002; 2002US-0373826P.
PR
19-APR-2002; 2002US-0373884P.
PR
22-APR-2002; 2002US-0374977P.
PR
16-MAY-2002; 2002US-0381037P.
PR
16-MAY-2002; 2002US-0381038P.
PR
16-MAY-2002; 2002US-0381042P.
PR
17-MAY-2002; 2002US-0381642P.
PR
28-MAY-2002; 2002US-0383656P.
PR
29-MAY-2002; 2002US-0383831P.
PR
25-JUN-2002; 2002US-0391335P.
PR
01-OCT-2002; 2002US-00262511.
(CURA-) CURAGEN CORP.

Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X; Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM; Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E; Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG; Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Bergins C, Dipippo VA; Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK; WPI; 2003-381626/36.
N-ESDB; ADA05731.

New NOVX polypeptides and nucleic acids, useful for diagnosing, preventing or treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or dyslipidemia, and in chromosome mapping, tissue typing or pharmacogenomics.

Claim 1; Page 169-170; 586pp; English.

The present invention describes NOVX proteins, where X can be 1 to 55 (e.g. NOV1). Also described: (1) a composition comprising a polypeptide described above and a carrier; (2) a kit comprising, in one or more containers, the composition described above; (3) an isolated nucleic acid

CC molecule which encodes a NOVX protein of the invention; (4) a vector
 CC comprising the nucleic acid molecule described above; (5) a cell
 CC comprising the above vector; (6) an antibody that immunospecifically
 CC binds to the polypeptide described above; (7) methods for determining the
 CC presence or amount of the above polypeptide or nucleic acid molecule in a
 CC sample; (8) methods for determining the presence of or predisposition to
 CC a disease associated with altered levels of expression of the above
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
 CC method of identifying an agent that binds to the polypeptide described
 CC above; (10) a method for identifying a potential therapeutic agent for
 CC use in treating a pathology that is related to an aberrant expression or
 CC aberrant physiological interactions of the polypeptide; (11) a method of
 CC screening for a modulator of activity or of latency or predisposition to
 CC a pathology associated with the polypeptide; (12) a method for modulating
 CC the activity of the polypeptide described above; (13) methods of treating
 CC or preventing a pathology associated with the above polypeptide in a
 CC mammal; and (14) a method for producing the above polypeptide. NOVX
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,
 CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian
 CC and antilipidemic activities, and can be used in gene therapy. The
 CC polypeptide is useful in manufacturing a medicament for treating a
 CC syndrome associated with a human disease. The polypeptide or the nucleic
 CC acid molecule may be used to diagnose, treat or prevent metabolic
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
 CC disease, immune disorders, haematopoietic disorders and various
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation
 CC probes, in chromosome mapping, tissue typing, preventive medicine and
 CC pharmacogenomics. The present sequence represents a human NOVX from the
 CC present invention.

XX Sequence 250 AA;

Query Match 100.0%; Score 49; DB 6; Length 250;
 Best Local Similarity 100.0%; Pred. NO. 0.24;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNEYTVHL 9

DB 69 KNEYTVHL 77

RESULT 11

ADA05734
 ID ADA05734 standard; protein; 252 AA.

AC ADA05734;

DT 06-NOV-2003 (first entry)

XX Human NOV18b protein SEQ ID NO:94.

XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
 KW immunomodulator; cytostatic; nootropic; neuroprotective;
 KW antiparkinsonian; antilipidemic; gene therapy; human disease;
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; dyslipidaemia.

XX Homo sapiens.

PN WO2003029424-A2.

XX 10-APR-2003.

XX 02-OCT-2002; 2002WO-US031373.

XX 02-OCT-2001; 2001US-0326483P.

PR 05-OCT-2001; 2001US-0327435P.

PR 05-OCT-2001; 2001US-0327443P.

PR 09-OCT-2001; 2001US-0327517P.

PR 09-OCT-2001; 2001US-0328029P.

PR 09-OCT-2001; 2001US-0328044P.

PR 09-OCT-2001; 2001US-0328056P.
 PR 12-OCT-2001; 2001US-0328049P.
 PR 15-OCT-2001; 2001US-0329141P.
 PR 17-OCT-2001; 2001US-0330142P.
 PR 18-OCT-2001; 2001US-0330309P.
 PR 22-OCT-2001; 2001US-0341058P.
 PR 24-OCT-2001; 2001US-0339266P.
 PR 24-OCT-2001; 2001US-0343629P.
 PR 29-OCT-2001; 2001US-0349575P.
 PR 01-NOV-2001; 2001US-0346357P.
 PR 17-APR-2002; 2002US-0373260P.
 PR 19-APR-2002; 2002US-0373815P.
 PR 19-APR-2002; 2002US-0373817P.
 PR 19-APR-2002; 2002US-0373826P.
 PR 19-APR-2002; 2002US-0373884P.
 PR 22-APR-2002; 2002US-0374977P.
 PR 16-MAY-2002; 2002US-0381037P.
 PR 16-MAY-2002; 2002US-0381038P.
 PR 16-MAY-2002; 2002US-0381042P.
 PR 17-MAY-2002; 2002US-0381642P.
 PR 28-MAY-2002; 2002US-0383656P.
 PR 29-MAY-2002; 2002US-0383831P.
 PR 25-JUN-2002; 2002US-0391335P.
 PR 01-OCT-2002; 2002US-00262511.

(CURA-) CURAGEN CORP.

XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
 PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
 PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Cattarton E;
 PI Ji W, Miller CE, Rastelli L, Stone DU, Pena CA, Shenoy SG;
 PI Shimkets RA, Rothenberg ME, Leach MD, Agee CA, Berghs C, Dippio VA;
 PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;

WPI: 2003-381626/36.

DR N-PSDB; ADA05733.

XX New NOVX polypeptides and nucleic acids, useful for diagnosing,
 PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
 PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.

XX Claim 1; Page 170; 586pp; English.

XX The present invention describes NOVX proteins, where X can be 1 to 55
 CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide
 CC described above and a carrier; (2) a kit comprising, in one or more
 CC containers, the composition described above; (3) an isolated nucleic acid
 CC molecule which encodes a NOVX protein of the invention; (4) a vector
 CC comprising the nucleic acid molecule described above; (5) a cell
 CC comprising the above vector; (6) an antibody that immunospecifically
 CC binds to the polypeptide described above; (7) methods for determining the
 CC presence or amount of the above polypeptide or nucleic acid molecule in a
 CC sample; (8) methods for determining the presence of or predisposition to
 CC a disease associated with altered levels of expression of the above
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
 CC method of identifying an agent that binds to the polypeptide described
 CC above; (10) a method for identifying a potential therapeutic agent for
 CC use in treating a pathology that is related to an aberrant expression or
 CC aberrant physiological interactions of the polypeptide; (11) a method of
 CC screening for a modulator of activity or of latency or predisposition to
 CC a pathology associated with the polypeptide; (12) a method for modulating
 CC the activity of the polypeptide described above; (13) methods of treating
 CC or preventing a pathology associated with the above polypeptide in a
 CC mammal; and (14) a method for producing the above polypeptide. NOVX
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,
 CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian
 CC and antilipidemic activities, and can be used in gene therapy. The
 CC polypeptide is useful in manufacturing a medicament for treating a
 CC syndrome associated with a human disease. The polypeptide or the nucleic
 CC acid molecule may be used to diagnose, treat or prevent metabolic
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's

CC disease, immune disorders, haematopoietic disorders and various
CC dyslipidaemias. The nucleic acids can also be used as hybridisation
CC probes, in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. The present sequence represents a human NOVX from the
CC present invention.

XX SQ Sequence 252 AA;

Query Match 100.0%; Score 49; DB 6; Length 252;

Best Local Similarity 100.0%; Pred. No. 0.24; Mismatches 0; Indels 0; Gaps 0;

QY 1 KMNEYTVHL 9

DB 71 KMNEYTVHL 79

RESULT 12

AAR67888

ID AAR67888 standard; protein; 253 AA.

XX AC AAR67888;

XX DT 25-MAR-2003 (revised)

XX DT 09-AUG-1995 (first entry)

XX DE Human stratum corneum chymotryptic recombinant enzyme (SCCE).

XX XX Stratum corneum chymotryptic enzyme; skin disorder; acne; psoriasis;

XX KW callosities; keratosis pilaris; ichthyoses; eczema.

XX OS Homo sapiens.

XX PN WO9500651-A1.

XX PD 05-JAN-1995.

XX PF 20-JUN-1994; 94WO-IB000166.

XX PR 18-JUN-1993; 93DK-00000725.

XX PA (SYMB-) SYMBICOM AB.

XX PI Egelrud T, Hansson L;

XX DR WPI; 1995-052088/07.

XX DR N-PSDB; AAQ81203.

XX PT Nucleotide sequences encoding stratum corneum chymotryptic enzyme - and

XX PT related vectors, transformed cells and polypeptides, useful for treating

XX PT skin disorders, e.g. acne or psoriasis, and for identification of

XX PT specific inhibitors.

XX PS Disclosure; Page 97; 137pp; English.

XX CC The enzyme encoded by this sequence is used in pharmaceutical, cosmetic

XX CC and skin care products, especially to treat and prevent acne, xeroderma,

XX CC or other hyperkeratotic conditions (e.g. callosities or keratosis

XX CC pilaris), ichthyoses, psoriasis, eczema, etc. It is produced

XX CC recombinantly following mammalian, insect, plant, or microorganism

XX CC transformation with plasmid pS507. (Updated on 25-MAR-2003 to correct PN

XX CC field.)

XX SQ Sequence 253 AA;

Query Match 100.0%; Score 49; DB 2; Length 253;

Best Local Similarity 100.0%; Pred. No. 0.24; Mismatches 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KMNEYTVHL 9

DB 72 KMNEYTVHL 80

Query Match 100.0%; Score 49; DB 2; Length 253;

Best Local Similarity 100.0%; Pred. No. 0.24; Mismatches 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KMNEYTVHL 9

DB 72 KMNEYTVHL 80

RESULT 13

AAW05383

ID AAW05383 standard; protein; 253 AA.

XX AC AAW05383;

XX DT 31-DEC-1996 (first entry)

XX DE Human amyloid precursor protein protease.

XX KW Amyloid precursor protein protease; Alzheimer's disease; diagnosis;

XX KW therapy.

XX OS Homo sapiens.

XX PN WO9631122-A1.

XX PD 10-OCT-1996.

XX PF 02-APR-1996; 96WO-US004294.

XX PR 04-APR-1995; 95US-00416257.

XX PA (ELIL) LILLY & CO ELI.

XX PI Dixon EP, Johnstone EM, Little SP;

XX DR WPI; 1996-464694/46.

XX DR N-PSDB; AAT39783.

XX PT New isolated human amyloid precursor protein protease - used to develop

XX PT prods. for the treatment or diagnosis of associated conditions, esp.

XX PT Alzheimer's disease.

XX PS Claim 1; Page 44-45; 55pp; English.

XX CC Human amyloid precursor protein protease (AAW05383) is involved in the

XX CC processing or clearance of amyloid precursor protein to form beta-amyloid

XX CC peptide. Its amino acid sequence was deduced from a cDNA clone (AAT39783)

XX CC obtd. from a human lung library. Recombinant protease can be produced in

XX CC transformed or transfected prokaryotic (partic. E. coli) or eukaryotic

XX CC (partic. AV-120 host cells. It is used to develop products for the design

XX CC and testing of cpds. useful for treating or preventing conditions

XX CC associated with beta-amyloid peptide, esp. Alzheimer's disease

XX SQ Sequence 253 AA;

Query Match 100.0%; Score 49; DB 2; Length 253;

Best Local Similarity 100.0%; Pred. No. 0.24; Mismatches 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KMNEYTVHL 9

DB 72 KMNEYTVHL 80

RESULT 14

ABB84421

ID ABB84421 standard; peptide; 253 AA.

XX AC ABB84421;

XX DT 08-NOV-2002 (first entry)

XX DE Human SCCE protein N-terminal fragment SEQ ID 48.

XX KW SCCE; human; stratum corneum chymotryptic enzyme; kallikrein 7;

XX KW serine protease; transgenic mammal; skin; skin disease; skin cancer;

XX KW hyperkeratosis; acanthosis; epidermal inflammation; dermal inflammation;

XX KW pruritus; atopic dermatitis; eczema; acne; itch; KLK7; enzyme.

XX OS Homo sapiens.

```

XX PN WO200262135-A2.
XX XX
XX PD 15-AUG-2002.
XX XX
XX PF 08-FEB-2002; 2002WO-IB001300.
XX XX
XX PR 09-FEB-2001; 2001CA-02332855.
XX PR 09-FEB-2001; 2001DK-00000218.
XX PA (EGEL/) EGELRUD T.
XX PA (HANS/) HANSSON L.
XX XX
XX PI Egelrud T, Hansson L;
XX DR WPI; 2002-643380/69.
XX XX
XX FT Transgenic mammal or its embryo useful as model for human disease, has
XX PT heterologous nucleotide sequence coding for stratum corneum chymotryptic
XX PT enzyme operably linked to promoter that drives its expression in skin.
XX XX
XX PS Example 6; Page 37; 74pp; English.
XX CC
XX CC This invention describes a novel non-human transgenic mammal or mammalian
XX CC embryo having integrated within its genome, a heterologous nucleotide
XX CC sequence comprising at least a significant part of a nucleotide sequence
XX CC coding for a stratum corneum chymotryptic enzyme (SCCE) or its variant.
XX CC operably linked to a promoter that drives expression of heterologous scce
XX CC or its variant in skin. The product of the invention is useful as a model
XX CC for the study of disease with the aim of improving treatment, to relieve
XX CC or ameliorate a pathogenic condition, for development or testing of a
XX CC cosmetic or a pharmaceutical formulation, and for the development of a
XX CC diagnostic method. It can also be used as a model for a skin disease or
XX CC skin cancer. The invention is also useful for screening or identifying a
XX CC compound or composition effective for the prevention or treatment of an
XX CC abnormal or unwanted phenotype, and for screening or identifying a
XX CC compound or composition effective for the prevention or treatment of
XX CC inflammatory skin diseases selected from diseases consisting of epidermal
XX CC hyperkeratosis, acanthosis, epidermal inflammation, dermal inflammation,
XX CC pruritus, atopic dermatitis, eczema, acne and inherited skin diseases
XX CC with epidermal hyperkeratosis. The mammal of the invention is also useful
XX CC as a model for further studies of itch mechanisms and the testing of
XX CC potential compounds and compositions for relieve of various skin diseases
XX CC where itch is a component. This sequence represents the N-terminal
XX CC fragment of the human stratum corneum chymotryptic enzyme, SCCE
XX CC synonymous with human kallikrein 7 (KLK7), used in the development of the
XX CC transgenic mammals described in the invention
XX XX
XX SQ Sequence 253 AA;
XX CC
XX CC Query Match 100.0%; Score 49; DB 5; Length 253;
XX CC Best Local Similarity 100.0%; Pred. No. 0.24;
XX CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX Qy 1 KKNYEYTVHL 9
XX Db |||||
XX 72 KKNYEYTVHL 80
XX
XX RESULT 15
XX ABB84406
XX ID ABB84406 standard; protein; 253 AA.
XX AC
XX AC ABB84406;
XX XX
XX DT 08-NOV-2002 (first entry)
XX XX
XX DE Human SCCE protein.
XX XX
XX KW SCCE; human; stratum corneum chymotryptic enzyme; kallikrein 7;
XX KW serine protease; transgenic mammal; skin; skin disease; skin cancer;
XX KW hyperkeratosis; acanthosis; epidermal inflammation; dermal inflammation;
XX KW pruritus; atopic dermatitis; eczema; acne; itch; KLK7; enzyme.

```

```

XX OS Homo sapiens.
XX XX
XX PN WO200262135-A2.
XX XX
XX PD 15-AUG-2002.
XX XX
XX PF 08-FEB-2002; 2002WO-IB001300.
XX XX
XX PR 09-FEB-2001; 2001CA-02332855.
XX PR 09-FEB-2001; 2001DK-00000218.
XX XX
XX PA (EGEL/) EGELRUD T.
XX PA (HANS/) HANSSON L.
XX XX
XX PI Egelrud T, Hansson L;
XX XX
XX DR WPI; 2002-643380/69.
XX DR N-PSDB; ABQ76226.
XX XX
XX PT Transgenic mammal or its embryo useful as model for human disease, has
XX PT heterologous nucleotide sequence coding for stratum corneum chymotryptic
XX PT enzyme operably linked to promoter that drives its expression in skin.
XX XX
XX PS Claim 10; Page 58-59; 74pp; English.
XX CC
XX CC This invention describes a novel non-human transgenic mammal or mammalian
XX CC embryo having integrated within its genome, a heterologous nucleotide
XX CC sequence comprising at least a significant part of a nucleotide sequence
XX CC coding for a stratum corneum chymotryptic enzyme (SCCE) or its variant,
XX CC operably linked to a promoter that drives expression of heterologous scce
XX CC or its variant in skin. The product of the invention is useful as a model
XX CC for the study of disease with the aim of improving treatment, to relieve
XX CC or ameliorate a pathogenic condition, for development or testing of a
XX CC cosmetic or a pharmaceutical formulation, and for the development of a
XX CC diagnostic method. It can also be used as a model for a skin disease or
XX CC skin cancer. The invention is also useful for screening or identifying a
XX CC compound or composition effective for the prevention or treatment of an
XX CC abnormal or unwanted phenotype, and for screening or identifying a
XX CC compound or composition effective for the prevention or treatment of
XX CC inflammatory skin diseases selected from diseases consisting of epidermal
XX CC hyperkeratosis, acanthosis, epidermal inflammation, dermal inflammation,
XX CC pruritus, atopic dermatitis, eczema, acne and inherited skin diseases
XX CC with epidermal hyperkeratosis. The mammal of the invention is also useful
XX CC as a model for further studies of itch mechanisms and the testing of
XX CC potential compounds and compositions for relieve of various skin diseases
XX CC where itch is a component. This sequence represents the human stratum
XX CC corneum chymotryptic enzyme, SCCE which is a serine protease synonymous
XX CC with human kallikrein 7 (KLK7) and is used in the development of the
XX CC transgenic mammals described in the invention
XX XX
XX SQ Sequence 253 AA;
XX CC
XX CC Query Match 100.0%; Score 49; DB 5; Length 253;
XX CC Best Local Similarity 100.0%; Pred. No. 0.24;
XX CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX Qy 1 KKNYEYTVHL 9
XX Db |||||
XX 72 KKNYEYTVHL 80
XX
XX Search completed: March 1, 2004, 17:28:46
XX Job time : 47.5556 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:21:01 ; Search time 10.2222 Seconds
(without alignments)
84.690 Million cell updates/sec

Title: US-09-905-083-31

Perfect score: 49

Sequence: 1 KMNEYTVHL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	100.0	253	2 A53968	serine proteinase
2	39	79.6	218	1 YVBPX3	lysis protein t -
3	39	79.6	218	1 YVBP74	lysis protein t -
4	37	75.5	334	2 B98019	conserved hypotet
5	37	75.5	393	2 D75207	hypothetical prote
6	36	73.5	627	2 T25395	hypothetical prote
7	36	73.5	1829	2 T34239	hypothetical prote
8	35	71.4	159	2 A11406	spermidine/spermin
9	35	71.4	159	2 A11782	spermidine/spermin
10	35	71.4	505	2 H75431	conserved hypotet
11	35	71.4	555	2 AD1794	acylase and dieste
12	35	71.4	667	2 T09482	ring finger protei
13	35	71.4	667	2 T09013	RING finger protei
14	34	69.4	270	2 A43711	replication protei
15	34	69.4	270	2 S28682	replication protei
16	34	69.4	285	2 F70346	hypothetical prote
17	34	69.4	421	2 AC1481	conserved hypotet
18	34	69.4	421	2 AH1120	B. subtilis YwbN P
19	33	67.3	204	2 B31227	nitrogen fixation
20	33	67.3	204	2 E95345	FixJ Transcription
21	33	67.3	218	2 S53354	calflagin Tb-24 -
22	33	67.3	218	2 E83724	hypothetical prote
23	33	67.3	227	2 D83796	two-component resp
24	33	67.3	229	2 S53355	calflagin Tb-1.7 -
25	33	67.3	233	1 AQU117	flagellar calcium-
26	33	67.3	407	2 S53353	calflagin Tb-44A -
27	33	67.3	422	2 T11714	hypothetical prote
28	33	67.3	429	2 H70307	preprotein translo
29	33	67.3	525	2 B84028	oligopeptide ABC t

30 33 67.3 664 2 C72379 hypothetical prote
31 33 67.3 677 2 A87470 TonB-dependent rec
32 33 67.3 1042 2 S23738 pyr1-3 protein - s
33 33 67.3 1308 2 E71622 probable membrane
34 33 67.3 1481 1 OZDOP3 pyrimidine synthase
35 32 65.3 49 2 H89908 50S ribosomal prot
36 32 65.3 84 2 A82172 hypothetical prote
37 32 65.3 116 2 S24989 ribosomal protein
38 32 65.3 130 2 E97262 hypothetical prote
39 32 65.3 144 2 F71215 hypothetical prote
40 32 65.3 192 2 G81301 MdaB protein homol
41 32 65.3 228 2 C90033 hypothetical prote
42 32 65.3 304 2 H75378 hypothetical prote
43 32 65.3 339 2 F97121 probable membrane
44 32 65.3 371 2 A89800 conserved hypotet
45 32 65.3 448 2 F82280 citrate/sodium sym

ALIGNMENTS

RESULT 1

A53968

serine proteinase SCCE precursor - human

N:Alternate names: stratum corneum chymotryptic enzyme

C:Species: Homo sapiens (man)

C>Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 22-Jun-1999

C/Accession: A53968

R/Hansson, L.; Stroemqvist, M.; Baeckman, A.; Wallbrandt, P.; Carlstein, A.; Egeirud, T.

J. Biol. Chem. 269, 19420-19426, 1994

A:Title: Cloning, expression, and characterization of stratum corneum chymotryptic enzyme

A:Reference number: A53968; MUID:94308225; PMID:8034709

A:Accession: A53968

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-253 <HAN>

A/Cross-references: GB:I33404; NID:G521214; PIDN:AAC37551.1; PID:G532504

C/Genetics:

A:Gene: GDB:PRSS6; SCCE

A/Cross-references: GDB:377730

A:Map position: 7q35-7q35

C/Superfamily: trypsin; trypsin homology

F:30-245/Domain: trypsin homology <TRY>

Query Match 100.0%; Score 49; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. NO. 0.024;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KMNEYTVHL 9

DB 72 KMNEYTVHL 80

RESULT 2

YVBPX3

lysis protein t - phase K3

C/Species: phase K3

A/Note: host Escherichia coli

C/Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 23-Jul-1999

C/Accession: A27083

R/Riede, I.

J. Bacteriol. 169, 2956-2961, 1987

A:Title: Lysis gene t of T-even bacteriophages: evidence that colicins and bacteriophage

A/Reference number: A27083; MUID:87250254; PMID:3597316

A/Accession: A27083

A:Molecule type: DNA

A:Residues: 1-218 <RIE>

A/Cross-references: GB:M16812; NID:G215503; PIDN:AAA88415.1; PID:G215504

A/Note: the author translated the codon CAA for residue 85 as Ile and CAG for residue 20;

C/Genetics:

A:Gene: t

C/Superfamily: phase T4 lysis protein t

C/Keywords: host cell lysis

Query Match 79.6%; Score 39; DB 1; Length 218;
 Best Local Similarity 87.5%; Pred. No. 2.2;
 Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 2 MNEYTVHL 9
 :|||||
 Db 139 MDEYTVHL 146

RESULT 3
 YVBFT4
 Lysis protein t - phage T4
 C:Species: phage T4
 A:Note: host Escherichia coli
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 23-Jul-1999
 C:Accession: JF0028; S07395
 R:Montag, D.; Degen, M.; Henning, U.
 Nucleic Acids Res. 15, 6736, 1987
 A:Title: Nucleotide sequence of gene t (lysis gene) of the E. coli phage T4.
 A:Reference number: S07395; MUID:87316934; PMID:3628006
 A:Accession: JF0028
 A:Molecule type: DNA
 A:Residues: 1-218 <MON>
 A:Cross-references: GB:Y00408; NID:G15369; PIDN:CAA68470.1; PID:G15369
 A:Note: the sequence is almost identical with that of the E. coli phage K3
 C:Comment: At the end of the growth cycle, phage T4 expresses two genes with lysis function about the gene product of t, although it has been suggested that it acts as a phospholipase C:Genetics:
 A:Gene: t
 A:Map position: 157,985-158,639
 C:Superfamily: phage T4 lysis protein t
 C:Keywords: host cell lysis; transmembrane protein
 F:35-49/Domain: transmembrane #status predicted <TN>

Query Match 79.6%; Score 39; DB 1; Length 218;
 Best Local Similarity 87.5%; Pred. No. 2.2;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MNEYTVHL 9
 :|||||
 Db 139 MDEYTVHL 146

RESULT 4
 B98019
 conserved hypothetical protein spr1179 [imported] - Streptococcus pneumoniae (strain R6)
 C:Species: Streptococcus pneumoniae
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
 C:Accession: B98019
 R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; Ee, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; Meyer, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
 A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A:Reference number: A97872; MUID:21423245; PMID:11544234
 A:Accession: B98019
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-334 <KUR>
 A:Cross-references: GB:AE007317; PIDN:AAK99982.1; PID:G15458811; GSPDB:GN00174
 C:Genetics:
 A:Gene: spr1179

Query Match 75.5%; Score 37; DB 2; Length 334;
 Best Local Similarity 66.7%; Pred. No. 8.9;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KNEYTVHL 9
 :|||||
 Db 302 KNEYTVHL 310

RESULT 5
 D75207
 hypothetical protein PAB2235 - Pyrococcus abyssi (strain Orsay)
 C:Species: Pyrococcus abyssi
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 28-Jul-2000
 C:Accession: D75207
 R:anonymous, Genoscope
 submitted to the EMBL Data Library, July 1999
 A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure
 A:Reference number: A75001
 A:Accession: D75207
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-393 <KAW>
 A:Cross-references: GB:AJ248283; GB:AL096836; NID:G5457433; PIDN:CA849107.1; PID:G5457616
 A:Experimental source: strain Orsay
 C:Genetics:
 A:Gene: PAB2235
 C:Superfamily: Pyrococcus horikoshii hypothetical protein PH0196

Query Match 75.5%; Score 37; DB 2; Length 393;
 Best Local Similarity 66.7%; Pred. No. 11;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KNEYTVHL 9
 :|||||
 Db 129 KPEYTVHL 137

RESULT 6
 T25395
 hypothetical protein T28A8.6 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
 C:Accession: T25395
 R:Ridley, C.
 submitted to the EMBL Data Library, March 1997
 A:Reference number: Z20027
 A:Accession: T25395
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-627 <WIL>
 A:Cross-references: EMBL:Z292813; PIDN:CAB07289.1; GSPDB:GN00021; CESP:T28A8.6
 A:Experimental source: Clone T28A8
 C:Genetics:
 A:Gene: CESP:T28A8.6
 A:Map position: 3
 A:Introns: 51/2; 89/2; 183/2; 221/2; 296/2; 607/3
 C:Superfamily: Caenorhabditis elegans hypothetical protein Y57A10A.d

Query Match 73.5%; Score 36; DB 2; Length 627;
 Best Local Similarity 56.7%; Pred. No. 28;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KNEYTVHL 9
 :|||||
 Db 62 KNEYSVIEL 70

RESULT 7
 T34239
 hypothetical protein F26F12.7 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T34239
 R:Wilson, R.; Bentley, D.; Gattung, S.
 submitted to the EMBL Data Library, April 1996
 A:Description: The sequence of C. elegans cosmid F26F12.
 A:Reference number: Z21493
 A:Accession: T34239
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A;Residues: 1-1829 <WIL>
 A;Cross-references: EMBL:U55373; PIDN:AC25894.1; GSPDB:GN00023; CESP:F26F12.7
 A;Experimental source: strain Bristol N2; clone F26F12
 C;Genetics:
 A;Gene: CESP:F26F12.7
 A;Map position: 5
 A;Introns: 110/3; 441/3; 801/2; 1244/3; 1693/2; 1784/1

Query Match 73.5%; Score 36; DB 2; Length 1829;
 Best Local Similarity 71.4%; Pred. No. 89;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 KMNEYTVH 8
 Db 765 LNEYTVH 771
 |||:|

RESULT 8
 A11406
 spermidine/spermine N1-acetyl transferase homolog lmo2658 [imported] - *Listeria monocytogenes*
 C;Species: *Listeria monocytogenes*
 C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
 C;Accession: A11406
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 A;Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Matok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of *Listeria* species.
 A;Reference number: AB1077; MUID:21537279; PMID:11679669
 A;Accession: A11406
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-159 <GLA>
 A;Cross-references: GB:NC_003210; PIDN:CAD00871.1; PID:g16412158; GSPDB:GN00177
 A;Experimental source: strain EGD-e
 C;Genetics:
 A;Gene: lmo2658

Query Match 71.4%; Score 35; DB 2; Length 159;
 Best Local Similarity 75.0%; Pred. No. 10;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KMNEYTVH 8
 Db 139 KMSEWTVH 146
 |||:|

RESULT 9
 A11782
 spermidine/spermine N1-acetyl transferase homolog lin2807 [imported] - *Listeria innocua*
 C;Species: *Listeria innocua*
 C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
 C;Accession: A11782
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 A;Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Matok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of *Listeria* species.
 A;Reference number: AB1077; MUID:21537279; PMID:11679669
 A;Accession: A11782
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-159 <GLA>
 A;Cross-references: GB:AL592022; PIDN:CAC98033.1; PID:g16415343; GSPDB:GN00178
 A;Experimental source: strain Clp11262
 C;Genetics:
 A;Gene: lin2807

Query Match 71.4%; Score 35; DB 2; Length 159;

Best Local Similarity 75.0%; Pred. No. 10;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KMNEYTVH 8
 Db 139 KMSEWTVH 146
 |||:|

RESULT 10
 H75431
 conserved hypothetical protein - *Deinococcus radiodurans* (strain R1)
 C;Species: *Deinococcus radiodurans*
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
 C;Accession: H75431
 R;White, O.; Bisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; I.M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A;Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
 A;Reference number: A75250; MUID:20036896; PMID:10567266
 A;Accession: H75431
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-505 <WHI>
 A;Cross-references: GB:AE001964; GB:AE000513; NID:g6458881; PIDN:AAF10722.1; PID:g6458881
 A;Experimental source: strain R1
 C;Genetics:
 A;Gene: DR1150
 A;Map position: 1
 C;Superfamily: conserved hypothetical protein b0835

Query Match 71.4%; Score 35; DB 2; Length 505;
 Best Local Similarity 66.7%; Pred. No. 36;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KMNEYTVH 9
 Db 57 QMNEYDTHL 65
 |||:|

RESULT 11
 AD1794
 acylase and diesterase homolog lin2898 [imported] - *Listeria innocua* (strain Clp11262)
 C;Species: *Listeria innocua*
 C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
 C;Accession: AD1794
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 A;Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Matok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of *Listeria* species.
 A;Reference number: AB1077; MUID:21537279; PMID:11679669
 A;Accession: AD1794
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-555 <GLA>
 A;Cross-references: GB:AL592022; PIDN:CAC98124.1; PID:g16415433; GSPDB:GN00178
 A;Experimental source: strain Clp11262
 C;Genetics:
 A;Gene: lin2898

Query Match 71.4%; Score 35; DB 2; Length 555;
 Best Local Similarity 66.7%; Pred. No. 39;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KMNEYTVH 9
 Db 483 KINEYDIDL 491
 |||:|

RESULT 12


```

T09482
ring finger protein Fxy - human
C:Species: Homo sapiens (man)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 02-Sep-2000
C:Accession: T09482
R:Perry, J.; Feather, S.; Smith, A.; Palmer, S.; Ashworth, A.
submitted to the EMBL Data Library, November 1997
A:Description: The human FXY maps to chromosome Xp22.3: Implications for evolution of th
A:Reference number: Z16687
A:Accession: T09482
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-667 <PR>
A:Cross-references: EMBL:AF035360; NID:G2827993; PID:G2827994
C:Genetics:
A:Gene: FXY
A:Map position: Xp22.3
C:Superfamily: RING finger homology
F:6-65/Domain: RING finger homology <RRN>

Query Match 71.4%; Score 35; DB 2; Length 667;
Best Local Similarity 75.0%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KMNEYTVH 8
Db 444 KQHHTVH 451

RESULT 13
T09013
RING finger protein Fxy - mouse
C:Species: Mus musculus (house mouse)
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 02-Sep-2000
C:Accession: T09013
R:Palmer, S.; Perry, J.; Kipling, D.; Ashworth, A.
Proc. Natl. Acad. Sci. U.S.A. 94, 12030-12035, 1997
A:Title: A gene spans the pseudautosomal boundary in mice.
A:Reference number: Z16531; MUID:98004518; PMID:9342357
A:Accession: T09013
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-667 <PAL>
A:Cross-references: EMBL:AF026565; NID:G2589222; PIDN:AA83986.1; PID:G2589223
C:Genetics:
A:Gene: Fxy
A:Map position: X; Y
C:Superfamily: RING finger homology
C:Keywords: zinc finger
F:6-65/Domain: RING finger homology <RRN>

Query Match 71.4%; Score 35; DB 2; Length 667;
Best Local Similarity 75.0%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KMNEYTVH 8
Db 444 KQHHTVH 451

RESULT 14
A43711
replication protein repA 32K chain - human
C:Species: Homo sapiens (man)
C>Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 16-Feb-2001
C:Accession: A43711
R:Erdfiele, L.F.; Wold, M.S.; Kelly, T.J.
J. Biol. Chem. 265, 3177-3182, 1990
A:Title: The primary structure of the 32-kDa subunit of human replication protein A.
A:Reference number: A43711; MUID:90153966; PMID:2406247
A:Accession: A43711
A:Status: preliminary
A:Molecule type: mRNA

```

```

A:Residues: 1-270 <ERD>
A:Cross-references: EMBL:J05249; NID:G337349; PIDN:AAA3560.1; PID:G337350
C:Superfamily: Schizosaccharomyces pombe single-stranded DNA binding protein 30K chain

Query Match 69.4%; Score 34; DB 2; Length 270;
Best Local Similarity 62.5%; Pred. No. 29;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 MNEYTVHL 9
Db 152 MNEFTTHI 159

RESULT 15
S28692
replication protein A 32K chain homolog - mouse
C:Species: Mus musculus (house mouse)
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 16-Feb-2001
C:Accession: S28692
R:Nakagawa, M.; Tsukada, S.; Soma, T.; Shimizu, Y.; Miyake, S.; Iwamatsu, A.; Sugiyama, I.
Nucleic Acids Res. 19, 4292, 1991
A:Title: cDNA cloning of the murine 30-kDa protein homologous to the 32-kDa subunit of hu
A:Reference number: S28692; MUID:91334145; PMID:1908076
A:Accession: S28692
A:Molecule type: mRNA
A:Residues: 1-270 <NAG>
A:Cross-references: EMBL:D00812; NID:G220583; PIDN:BAA00693.1; PID:G220584
C:Superfamily: Schizosaccharomyces pombe single-stranded DNA binding protein 30K chain

```

```

Query Match 69.4%; Score 34; DB 2; Length 270;
Best Local Similarity 62.5%; Pred. No. 29;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 MNEYTVHL 9
Db 152 MNEFTTHI 159

Search completed: March 1, 2004, 17:36:24
Job time : 12.3333 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:35:01 ; Search time 24.1111 Seconds
(without alignments)
78.818 Million cell updates/sec

Title: US-09-905-083-31

Perfect score: 49

Sequence: 1 KNEYTVHL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 21113259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/prodata/1/pubaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/1/pubaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/prodata/1/pubaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/prodata/1/pubaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/1/pubaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/prodata/1/pubaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/prodata/1/pubaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/prodata/1/pubaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/prodata/1/pubaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/prodata/1/pubaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/prodata/1/pubaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/prodata/1/pubaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/prodata/1/pubaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/prodata/1/pubaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/prodata/1/pubaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/prodata/1/pubaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/prodata/1/pubaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/prodata/1/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	100.0	9	9	US-09-918-243-31
2	49	100.0	9	9	US-09-918-243-89
3	49	100.0	9	9	US-09-918-243-109
4	49	100.0	9	9	US-09-905-083-31
5	49	100.0	9	9	US-09-905-083-89
6	49	100.0	9	9	US-09-905-083-109
7	49	100.0	144	14	US-09-796-294-4
8	49	100.0	144	14	US-10-461-787-4
9	49	100.0	253	9	US-09-888-615-98
10	49	100.0	253	9	US-09-764-762-3
11	49	100.0	253	14	US-10-264-283-90
12	49	100.0	253	15	US-10-295-027-498
13	49	100.0	253	15	US-10-173-999-48
14	39	79.6	9	9	US-09-918-243-122
15	39	79.6	9	9	US-09-905-083-122

```

16 37 75.5 9 9 US-09-918-243-93
17 37 75.5 9 9 US-09-905-083-93
18 37 75.5 334 9 US-09-815-242-13249
19 35 71.4 64 14 US-10-029-386-28681
20 34 69.4 270 14 US-10-300-453A-43
21 34 69.4 341 15 US-10-334-143-74
22 33 67.3 360 14 US-10-156-761-11255
23 32 65.3 382 14 US-10-128-714-8374
24 32 65.3 398 9 US-09-815-242-5469
25 32 65.3 409 9 US-09-815-242-12418
26 32 65.3 446 10 US-09-955-999-67
27 32 65.3 528 15 US-10-374-780A-2112
28 32 65.3 688 14 US-10-038-010-50
29 32 65.3 752 9 US-09-815-242-13421
30 32 65.3 752 9 US-09-815-242-13661
31 31 63.3 85 9 US-09-864-761-39492
32 31 63.3 85 9 US-10-094-749-2338
33 31 63.3 178 15 US-10-104-047-3741
34 31 63.3 178 15 US-10-108-260A-3572
35 31 63.3 526 14 US-10-032-201B-312
36 31 63.3 575 9 US-09-839-136-2
37 31 63.3 711 9 US-09-738-626-4090
38 31 63.3 1084 13 US-10-071-900-3
39 31 63.3 1183 9 US-09-870-759-45
40 31 63.3 1183 10 US-09-751-708A-45
41 30 61.2 159 15 US-10-369-493-8956
42 30 61.2 166 9 US-09-815-242-13009
43 30 61.2 207 15 US-10-379-127-24
44 30 61.2 276 9 US-09-815-242-5843
45 30 61.2 342 14 US-10-238-075-730

```

ALIGNMENTS

RESULT 1

```

US-09-918-243-31
; Sequence 31, Application US/09918243
; Patent No. US20020142317A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D62230IP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 31
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 72-80 of the SCCE protein
US-09-918-243-31

```

```

Query Match 100.0%; Score 49; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 KNEYTVHL 9
    |||||
Db 1 KNEYTVHL 9

```

RESULT 2

```

US-09-918-243-89
; Sequence 89, Application US/09918243
; Patent No. US20020142317A1
; GENERAL INFORMATION:

```

; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 89
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 72-80 of the SCCE protein
US-09-918-243-89

Query Match 100.0%; Score 49; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KXNEYTVHL 9
Db 1 KXNEYTVHL 9

RESULT 3

US-09-918-243-109
; Sequence 109, Application US/09918243
; Patent No. US20020142317A1
; GENERAL INFORMATION:

; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 109
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 72-80 of the SCCE protein
US-09-918-243-109

Query Match 100.0%; Score 49; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KXNEYTVHL 9
Db 1 KXNEYTVHL 9

RESULT 4

US-09-905-083-31
; Sequence 31, Application US/09905083
; Patent No. US20020146708A1
; GENERAL INFORMATION:

; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP/C/Div
; CURRENT APPLICATION NUMBER: US/09/905,083
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 109
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 72-80 of the SCCE protein
US-09-905-083-31

; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 31
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 72-80 of the SCCE protein
US-09-905-083-31

Query Match 100.0%; Score 49; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KXNEYTVHL 9
Db 1 KXNEYTVHL 9

RESULT 5

US-09-905-083-89
; Sequence 89, Application US/09905083
; Patent No. US20020146708A1
; GENERAL INFORMATION:

; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP/C/Div
; CURRENT APPLICATION NUMBER: US/09/905,083
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 09/502,600
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 89
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 72-80 of the SCCE protein
US-09-905-083-89

Query Match 100.0%; Score 49; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KXNEYTVHL 9
Db 1 KXNEYTVHL 9

RESULT 6

US-09-905-083-109
; Sequence 109, Application US/09905083
; Patent No. US20020146708A1
; GENERAL INFORMATION:

; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP/C/Div
; CURRENT APPLICATION NUMBER: US/09/905,083
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 09/502,600
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 109
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 72-80 of the SCCE protein
US-09-905-083-109

US-09-905-083-109

Query Match 100.0%; Score 49; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNEYTVHL 9
| | | | |
DB 1 KNEYTVHL 9

RESULT 7

US-09-796-294-4
; Sequence 4, Application US/09796294
; Patent No. US20020037581A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Underwood, Lowell J.
; TITLE OF INVENTION: Extracellular Serine Protease
; FILE REFERENCE: D6020CIP3
; CURRENT APPLICATION NUMBER: US/09/796,294
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 09/618,259
; PRIOR FILING DATE: 2000-07-18
; NUMBER OF SEQ ID NOS: 72
; SEQ ID NO 4
; LENGTH: 144
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; NAME/KEY: DOMAIN
; OTHER INFORMATION: Amino acid sequence of stratum corneum chymotryptic
; OTHER INFORMATION: enzyme (sccc) catalytic domain
US-09-796-294-4

Query Match 100.0%; Score 49; DB 9; Length 144;
Best Local Similarity 100.0%; Pred. No. 0.086;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNEYTVHL 9
| | | | |
DB 9 KNEYTVHL 17

RESULT 8

US-10-461-787-4
; Sequence 4, Application US/10461787
; Publication No. US20030199010A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Underwood, Lowell J.
; TITLE OF INVENTION: No. US20030199010A1e1 Extracellular Serine Protease
; FILE REFERENCE: D6020CIP2
; CURRENT APPLICATION NUMBER: US/10/461,787
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US/09/618,259
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: US 09/127,444
; PRIOR FILING DATE: 1998-08-21
; NUMBER OF SEQ ID NOS: 72
; SEQ ID NO 4
; LENGTH: 144
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; NAME/KEY: DOMAIN
; OTHER INFORMATION: Amino acid sequence of stratum corneum chymotryptic
; OTHER INFORMATION: enzyme (sccc) catalytic domain
US-10-461-787-4

Query Match 100.0%; Score 49; DB 14; Length 144;
Best Local Similarity 100.0%; Pred. No. 0.086;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNEYTVHL 9
| | | | |
DB 9 KNEYTVHL 17

RESULT 9

US-09-888-615-98
; Sequence 98, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENRPEEL, SEAN
; APPLICANT: CHARVDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 98
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-98

Query Match 100.0%; Score 49; DB 9; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNEYTVHL 9
| | | | |
DB 72 KNEYTVHL 80

RESULT 10

US-09-764-762-3
; Sequence 3, Application US/09764762
; Patent No. US20020068341A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NOVEL KALLIKREIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/764,762
; FILING DATE: 16-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/210,084
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0252 US
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 253 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: GenBank
 CLONE: 532504
 SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Query Match 100.0%; Score 49; DB 9; Length 253;
 Best Local Similarity 100.0%; Pred. NO. 0.15;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KXNEYTVHL 9
 |||||
 Db 72 KXNEYTVHL 80

RESULT 11

US-10-264-283-90
 Sequence 90, Application US/10264283
 Publication No. US2003014494A1
 GENERAL INFORMATION:
 APPLICANT: Algate, Paul A.
 APPLICANT: Manion, Jane
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
 FILE REFERENCE: 210121.590
 CURRENT APPLICATION NUMBER: US/10/264,283
 CURRENT FILING DATE: 2002-10-02
 NUMBER OF SEQ ID NOS: 111
 SOFTWARE: Corixa Invention Disclosure Database
 SEQ ID NO 90
 LENGTH: 253
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-264-283-90

Query Match 100.0%; Score 49; DB 14; Length 253;
 Best Local Similarity 100.0%; Pred. NO. 0.15;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KXNEYTVHL 9
 |||||
 Db 72 KXNEYTVHL 80

RESULT 12

US-10-295-027-498
 Sequence 498, Application US/10295027
 Publication No. US20030232350A1
 GENERAL INFORMATION:
 APPLICANT: Afar, Daniel
 APPLICANT: Aziz, Natasha
 APPLICANT: Ginsberg, Wendy M.
 APPLICANT: Gish, Kurt C.
 APPLICANT: Glynn, Richard
 APPLICANT: Hevezi, Peter A.
 APPLICANT: Mack, David H.
 APPLICANT: Murray, Richard
 APPLICANT: Watson, Susan R.
 APPLICANT: Eos Biotechnology, Inc.
 TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
 TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
 FILE REFERENCE: 018501-012500US
 CURRENT APPLICATION NUMBER: US/10/295,027
 CURRENT FILING DATE: 2002-11-13
 PRIOR APPLICATION NUMBER: US 09/663,733

PRIOR FILING DATE: 2000-09-15
 PRIOR APPLICATION NUMBER: US 60/350,666
 PRIOR FILING DATE: 2001-11-13
 PRIOR APPLICATION NUMBER: US 60/335,394
 PRIOR FILING DATE: 2001-11-15
 PRIOR APPLICATION NUMBER: US 60/332,464
 PRIOR FILING DATE: 2001-11-21
 PRIOR APPLICATION NUMBER: US 60/334,393
 PRIOR FILING DATE: 2001-11-29
 PRIOR APPLICATION NUMBER: US 60/340,376
 PRIOR FILING DATE: 2001-12-14
 PRIOR APPLICATION NUMBER: US 60/347,211
 PRIOR FILING DATE: 2002-01-08
 PRIOR APPLICATION NUMBER: US 60/347,349
 PRIOR FILING DATE: 2002-01-10
 PRIOR APPLICATION NUMBER: US 60/355,250
 PRIOR FILING DATE: 2002-02-08
 PRIOR APPLICATION NUMBER: US 60/356,714
 PRIOR FILING DATE: 2002-02-13
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 1386
 SOFTWARE: Patent In Ver. 2.1
 SEQ ID NO 498
 LENGTH: 253
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-295-027-498

Query Match 100.0%; Score 49; DB 15; Length 253;
 Best Local Similarity 100.0%; Pred. NO. 0.15;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KXNEYTVHL 9
 |||||
 Db 72 KXNEYTVHL 80

RESULT 13

US-10-173-999-48
 Sequence 48, Application US/10173999
 Publication No. US20040005563A1
 GENERAL INFORMATION:
 APPLICANT: Mack, David H.
 APPLICANT: Gish, Kurt C.
 APPLICANT: Eos Biotechnology, Inc.
 TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
 TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
 TITLE OF INVENTION: Cancer
 FILE REFERENCE: 018501-002420US
 CURRENT APPLICATION NUMBER: US/10/173,999
 CURRENT FILING DATE: 2002-06-17
 PRIOR APPLICATION NUMBER: US 60/299,234
 PRIOR FILING DATE: 2001-06-18
 PRIOR APPLICATION NUMBER: US 60/315,287
 PRIOR FILING DATE: 2001-08-27
 PRIOR APPLICATION NUMBER: US 60/350,666
 PRIOR FILING DATE: 2001-11-13
 PRIOR APPLICATION NUMBER: US 60/372,246
 PRIOR FILING DATE: 2001-04-12
 NUMBER OF SEQ ID NOS: 163
 SOFTWARE: Patent In Ver. 2.1
 SEQ ID NO 48
 LENGTH: 253
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-173-999-48

Query Match 100.0%; Score 49; DB 15; Length 253;
 Best Local Similarity 100.0%; Pred. NO. 0.15;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KXNEYTVHL 9
 |||||

Db 72 KMNEYTVHL 80

RESULT 14

US-09-918-243-122
; Sequence 122, Application US/09918243
; Patent No. US20020142317A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 122
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 74-82 of the SCCE protein
US-09-918-243-122

Query Match 79.6%; Score 39; DB 9; Length 9;
Best Local Similarity 100.0%; Pred.No. 7.1e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NEYTVHL 9
| | | | |
Db 1 NEYTVHL 7

RESULT 15

US-09-905-083-122
; Sequence 122, Application US/09905083
; Patent No. US20020146708A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP/C/Div
; CURRENT APPLICATION NUMBER: US/09/905,083
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 09/502,600
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 122
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 74-82 of the SCCE protein
US-09-905-083-122

Query Match 79.6%; Score 39; DB 9; Length 9;
Best Local Similarity 100.0%; Pred.No. 7.1e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NEYTVHL 9
| | | | |
Db 1 NEYTVHL 7

Search completed: March 1, 2004, 18:08:48
Job time : 25.1111 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:17:25 ; Search time 6 Seconds
(without alignments)
78.105 Million cell updates/sec

Title: US-09-905-083-31
Perfect score: 49
Sequence: 1 KMNEYTVHL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	100.0	253	1 KLK7 HUMAN	P49862 homo sapien
2	39	79.6	218	1 VLYS BFK3	P10393 bacterioph
3	39	79.6	218	1 VLYS BPT4	P06808 bacterioph
4	38	77.6	261	1 RFA4 HUMAN	Q13156 homo sapien
5	37	75.5	293	1 PRIL PYRAB	Q9V291 pyrococcus
6	35	71.4	667	1 MID1 HUMAN	Q15344 homo sapien
7	35	71.4	667	1 MID1 MUSSP	P82457 mus spratus
8	35	71.4	667	1 MID1 RAT	P82458 rattus norv
9	35	71.4	680	1 MID1 MOUSE	Q70583 mus musculus
10	35	71.4	685	1 MID2 MOUSE	Q9U1V3 homo sapien
11	35	71.4	715	1 MID2 HUMAN	P15927 homo sapien
12	34	69.4	266	1 RFA2 RAT	Q63528 rattus norv
13	34	69.4	270	1 RFA2 MOUSE	Q62193 mus musculus
14	34	69.4	285	1 Y539 AQUAE	O68818 aquifex aeo
15	34	69.4	285	1 Y539 RHIZO	P10958 rhizobium m
16	33	67.3	204	1 FIXJ RHIME	Q26680 trypanosoma
17	33	67.3	218	1 FCA2 TRYBB	Q26677 trypanosoma
18	33	67.3	229	1 FCA4 TRYBB	P17882 trypanosoma
19	33	67.3	231	1 FCAL TRYBB	P20666 moraxella b
20	33	67.3	243	1 TPFE MORBO	Q26678 trypanosoma
21	33	67.3	407	1 FCA2 TRYBB	Q6491 aquifex aeo
22	33	67.3	429	1 SECY AQUAE	Q92268 mus musculus
23	33	67.3	799	1 RSG4 MOUSE	O95294 homo sapien
24	33	67.3	804	1 RSG4 HUMAN	P20054 dictyosteli
25	33	67.3	2185	1 PYR1 DICDI	Q99U61 staphylococ
26	32	65.3	49	1 R332 STAAE	Q8CSP9 staphylococ
27	32	65.3	49	1 R332 CHLRE	P45941 chlamydomon
28	32	65.3	116	1 RL33 CHLRE	Q99IC7 staphylococ
29	32	65.3	228	1 RPIA STAAE	Q97157 clostridium
30	32	65.3	339	1 YH96 CLOAB	P19535 bos taurus
31	32	65.3	491	1 CAD3 BOVIN	P36906 thermoanaer
32	32	65.3	500	1 XYNB THESA	O30360 thermoanaer
33	32	65.3	500	1 XYNB_THESJ	

34	32	65.3	517	1 MATK ACEPS	Q8se90 acer pseudo
35	32	65.3	639	1 TET5 ENTPA	P11131 enterococcu
36	32	65.3	639	1 TETM STRPN	Q54807 streptococc
37	32	65.3	688	1 ARK2 HUMAN	P35626 homo sapien
38	32	65.3	752	1 CLPE STRPN	P35594 streptococc
39	32	65.3	987	1 YD94 METUA	O58789 methanococc
40	32	65.3	1127	1 Y855 TREPA	O83827 treponema p
41	31	63.3	228	1 RPE HUMAN	Q96at9 homo sapien
42	31	63.3	228	1 RPE MOUSE	Q8vee0 mus musculu
43	31	63.3	276	1 DAPF VIBCH	Q9Jvt5 vibrio chol
44	31	63.3	306	1 TRUB NEIMA	Q9Jjy1 neisseria m
45	31	63.3	307	1 TRUB_NEIMB	

ALIGNMENTS

RESULT 1

KLK7_HUMAN	STANDARD;	PRT;	253 AA.
ID_KLK7_HUMAN			
AC_P49862;			
DT_01-OCT-1996 (Rel. 34, Created)			
DT_01-OCT-1996 (Rel. 34, Last sequence update)			
DT_10-OCT-2003 (Rel. 42, Last annotation update)			
DE_Kallikrein 7 precursor (EC 3.4.21.-) (Stratum corneum chymotryptic enzyme) (hSCCE)			
GN_KLK7 OR PRS86 OR SCCE.			
OS_Homo sapiens (Human)			
OC_Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX_NCBI_TaxID=9606;			
RN	[1]		SEQUENCE FROM N.A., AND SEQUENCE OF 23-53.
RP	TISSUE=Skin.		
RC	TISSUE=Skin.		
RX	MEDLINE=94308225; PubMed=8034709;		
RA	Hansson L., Stroemqvist M., Baeckman A., Wallbrandt P., Carlstein A., Egelrud T.;		
RT	"Cloning, expression, and characterization of stratum corneum chymotryptic enzyme. A skin-specific human serine proteinase.";		
RL	J. Biol. Chem. 269:19420-19426(1994).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Keratinocytes;		
RA	Yousef G.M., Scorilas A., Diamandis E.P.;		
RT	"Molecular characterization, mapping and tissue expression of the human stratum corneum chymotryptic enzyme gene.";		
RL	Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=20510030; PubMed=11054574;		
RA	Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J., Moss P., Paepers B., Wang K.;		
RT	"Sequencing and expression analysis of the serine protease gene cluster located in chromosome 19q13 region.";		
RL	Gene 257:119-130(2000).		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RA	Hansson L., Baeckman A., Ny A., Edlund M., Edholm E., Tornell J., Wallbrandt P., Egelrud T.;		
RT	"Epidermal overexpression of stratum corneum chymotryptic enzyme in mice: a model for chronic itchy dermatitis.";		
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.		
RN	[5]		
RP	CHARACTERIZATION.		
RX	MEDLINE=95314630; PubMed=7794273;		
RA	Skytt A., Stroemqvist M., Egelrud T.;		
RT	"Primary substrate specificity of recombinant human stratum corneum chymotryptic enzyme.";		
RL	Biochem. Biophys. Res. Commun. 211:586-589(1995).		
CC	FUNCTION: May catalyze the degradation of intercellular cohesive structures in the cornified layer of the skin in the continuous shedding of cells from the skin surface. Specific for amino acid residues with aromatic side chains in the P1 position. SCCE		

```

CC cleaves insulin B chain at 6-Leu-|-Cys-7, 16-Tyr-|-Leu-17, 25-Phe-
CC -|-Tyr-26, and 26-Tyr-|-Thr-27. cytokines.
CC of precursors to inflammatory cytokines.
CC -|- TISSUE SPECIFICITY: Abundantly expressed in the skin and is
CC expressed by keratinocytes in the epidermis. Very low levels are
CC also seen in the brain and kidney.
CC -|- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L33404; AAC37551.1; --
CC EMBL: AF166330; AA049718.1; --
CC EMBL: AF243527; AAG33360.1; --
CC EMBL: AF332583; AAK69624.1; --
CC PIR: A53968; A53968.
CC HSSP: P00763; IDPO.
CC MEROPS: S01.300; --
CC Genew: HGNC:6368; KIK7.
CC MIN: 50438; --
CC GO: GO:0008236; F:serine-type peptidase activity; TAS.
CC GO: GO:0008544; P:epidermal differentiation; TAS.
CC InterPro: IPR009003; Cys Ser trypsin.
CC InterPro: IPR001254; Peptidase_S1.
CC InterPro: IPR001314; Peptidase_S1A.
CC Pfam: PF00089; trypsin.1
CC PRINTS: PR00722; CHYMOTRYPSIN.
CC SMART: SM00020; TRYPSIN_DOM.
CC PROSITE: PS0240; TRYPSIN_DOM; 1.
CC PROSITE: PS00134; TRYPSIN_HIS; 1.
CC PROSITE: PS00135; TRYPSIN_SER; 1.
CC KEGG: K00001; Serine protease; Zymogen; Glycoprotein; Signal.
CC
CC SIGNAL 1 22
CC PROPEP 23 29 ACTIVATION PEPTIDE.
CC CHAIN 30 253 KALLIKREIN 7.
CC ACT_SITE 70 70 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC ACT_SITE 112 112 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC ACT_SITE 205 205 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC DISULFID 36 137 BY SIMILARITY.
CC DISULFID 55 71 BY SIMILARITY.
CC DISULFID ? 239 BY SIMILARITY.
CC DISULFID 144 211 BY SIMILARITY.
CC DISULFID 176 190 BY SIMILARITY.
CC DISULFID 201 226 BY SIMILARITY.
CC CARBOHYD 246 246 N-LINKED (GLCNAC... ) (POTENTIAL).
CC SEQUENCE 253 AA; 27525 MW; 2D68B6B15A76A668 CRC64;
CC
CC Query Match 100.0%; Score 49; DB 1; Length 253;
CC Best Local Similarity 100.0%; Pred. NO. 0.012;
CC Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 KNEYTVVHL 9
CC DB 72 KNEYTVVHL 80
CC
CC RESULT 2
CC VLYS_BPK3 STANDARD; PRT; 218 AA.
CC AC P10393;
CC DT 01-MAR-1989 (Rel. 10, Created)
CC DT 01-MAR-1989 (Rel. 10, Last sequence update)
CC DT 01-JAN-1990 (Rel. 13, Last annotation update)
CC DE Lysis protein.
CC GN T.
CC OS Bacteriophage X3.
CC OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
CC OX NCBI_TaxID=10674;

```

```

RN SEQUENCE FROM N.A.
RX MEDLINE=87250254; PubMed=3597316;
RA Riede I.;
RT "Lysis gene t of T-even bacteriophages: evidence that colicins and
RT bacteriophage genes have common ancestors.";
RL J. Bacteriol. 169:2956-2961(1987).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M16812; AAA88415.1; --
CC PIR: A27083; IYBPK3.
CC KW Phage lysis protein.
CC SQ SEQUENCE 218 AA; 25222 MW; 21B4DC02ACA0ECF6 CRC64;
CC
CC Query Match 79.6%; Score 39; DB 1; Length 218;
CC Best Local Similarity 87.5%; Pred. NO. 1.1;
CC Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 2 KNEYTVVHL 9
CC DB 139 MDEYTVHL 146
CC
CC RESULT 3
CC VLYS_BPT4 STANDARD; PRT; 218 AA.
CC ID VLYS_BPT4
CC AC P06808;
CC DT 01-JAN-1988 (Rel. 06, Created)
CC DT 01-JAN-1988 (Rel. 06, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Lysis protein (Holin) (Protein rv).
CC T OR RV.
CC OS Bacteriophage T4.
CC OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
CC T4-like viruses.
CC OX NCBI_TaxID=10665;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=87316934; PubMed=3628006;
CC RA Montag D., Degen M., Henning U.;
CC RT "Nucleotide sequence of gene t (lysis gene) of the E. coli phage T4.";
CC RL Nucleic Acids Res. 15:6736-6736(1987).
CC RN [2]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=22514363; PubMed=12626685;
CC RA Miller E.S., Kutter E., Mosig G., Arisaka F., Kunisawa T., Ruger W.;
CC RT "Bacteriophage T4 genome.";
CC RL Microbiol. Mol. Biol. Rev. 67:86-156(2003).
CC RN [3]
CC RP SEQUENCE OF 1-34 FROM N.A.
CC RX MEDLINE=88011316; PubMed=2958637;
CC RA Montag D., Riede I., Eschbach M.-L., Degen M., Henning U.;
CC RT "Receptor-recognizing proteins of T-even type bacteriophages.
CC Constant and hypervariable regions and an unusual case of
CC evolution.";
CC RL J. Mol. Biol. 196:165-174(1987).
CC RN [4]
CC RP SEQUENCE OF 202-218 FROM N.A.
CC RC STRAIN=D;
CC RX MEDLINE=93106978; PubMed=8416914;
CC RA Orsini G., Ouhamouch M., le Caer J.P., Brody E.N.;
CC RT "The asiA gene of bacteriophage T4 codes for the anti-sigma 70
CC protein.";
CC RL J. Bacteriol. 175:85-93(1993).
CC CC -|- FUNCTION: At the end of the growth cycle, phage T4 expresses two
CC genes with lysis function, e and t. Nothing is known about the

```


CC gene product of t, although it has been suggested that it acts as
 CC a phospholipase.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Y00408; CAA68470.1; -;
 CC EMBL; AF158101; AAD42661.1; -;
 CC EMBL; X05677; CAA29164.1; -;
 CC EMBL; M99441; AAA32481.1; -;
 CC PIR; JF0028; YVBPT4.
 CC Phage lysis protein; 49
 CC TRANSMEM 35
 CC SEQUENCE 218 AA; 25175 MW; 9110BE11D772DF5 CRC64;
 CC
 CC Query Match 79.6%; Score 39; DB 1; Length 218;
 CC Best Local Similarity 87.5%; Pred. No. 1.1;
 CC Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 2 MNEYTVHL 9
 CC Db 139 MDEYTVHL 146
 CC
 CC RESULT 4
 CC RFA4 HUMAN
 CC ID_RFA4_HUMAN STANDARD; PRT; 261 AA.
 CC AC Q13156; -;
 CC DT 01-NOV-1997 (Rel. 35, Created)
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Replication protein A 30 kDa subunit (RP-A) (RP-A) (Replication
 CC factor-A protein 4).
 CC GN RPA4
 CC OS Homo sapiens (Human).
 CC OC Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CC OX NCBI_TaxID=9606;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=95280910; PubMed=7760808;
 CC RA Keshav K.F., Chen C., Dutta A.;
 CC RT "Rpa4, a homolog of the 34-kilodalton subunit of the replication
 CC protein A complex.";
 CC RL Mol. Cell. Biol. 15:3119-3128(1995).
 CC RN [2]
 CC RP REVISION TO 10.
 CC RA Keshav K.F.;
 CC RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 CC RN [3]
 CC RP SEQUENCE FROM N.A.
 CC RA Rieder M.J., Armel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,
 CC Ra Cuna M., Poel C.L., Toth E.J., Yi Q., Nickerson D.A.;
 CC RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 CC CC -!- FUNCTION: Absolutely required for simian virus 40 DNA replication
 CC in vitro. It participates in a very early step in initiation. RP-A
 CC is a single-stranded DNA-binding protein (By similarity).
 CC -!- SUBUNIT: Heterotrimer of 70, 32/30, and 14 kDa chains. The DNA-
 CC binding activity may reside exclusively on the 70 kDa subunit.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC CC -!- TISSUE SPECIFICITY: Preferentially expressed in placental and
 CC colon mucosa.
 CC CC -!- SIMILARITY: TO OTHER SPECIES RPA2/RPA2.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U24186; AAB08488.2; -;
 CC EMBL; AF494047; AAM09569.1; -;
 CC GK; Q13156; -;
 CC GO; GO:0005662; C:DNA replication factor A complex; TAS.
 CC GO; GO:0003697; F:single-stranded DNA binding; TAS.
 CC GO; GO:0006270; P:DNA replication initiation; TAS.
 CC InterPro; IPR008994; Nucleic acid OB.
 CC InterPro; IPR004365; tRNA anti; 1.
 CC Pfam; PF01336; tRNA anti; 1.
 CC DNA replication; Nuclear protein; Alternative splicing.
 CC SEQUENCE 261 AA; 28868 MW; 6A925FAEDBE21718 CRC64;
 CC
 CC Query Match 77.6%; Score 38; DB 1; Length 261;
 CC Best Local Similarity 75.0%; Pred. No. 2.1;
 CC Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 2 MNEYTVHL 9
 CC Db 151 MNEYTVHL 158
 CC
 CC RESULT 5
 CC PRIL PYRAB
 CC ID_PRIL_PYRAB STANDARD; PRT; 393 AA.
 CC AC Q9V291;
 CC DT 10-OCT-2003 (Rel. 42, Created)
 CC DT 10-OCT-2003 (Rel. 42, Last sequence update)
 CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
 CC DE DNA primase large subunit (EC 2.7.7.-) (DNA primase 46 kDa subunit)
 CC (p46).
 CC GN PRIB OR PYRAB01830 OR PAB2235.
 CC OS Pyrococcus abyssi.
 CC OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 CC OC Pyrococcus.
 CC OX NCBI_TaxID=29292;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX STRAIN=GES / Orsay;
 CC RX MEDLINE=22511545; PubMed=12622808;
 CC RA Cohen G.N., Barbe V., Flament D., Galperin M., Heilig R., Lecompte O.,
 CC Poch O., Prieur D., Querellou J., Ripp R., Thierry J.-C.,
 CC Van der Oost J., Weissenbach J., Zivanovic Y., Forterre P.;
 CC RT "An integrated analysis of the genome of the hyperthermophilic
 CC archaeon Pyrococcus abyssi.";
 CC RL Mol. Microbiol. 47:1495-1512(2003).
 CC CC -!- FUNCTION: DNA primase is the polymerase that synthesizes small RNA
 CC primers for the Okazaki fragments on both template strands at
 CC replication forks during chromosomal DNA synthesis (By
 CC similarity).
 CC CC -!- SUBUNIT: Heterodimer of a small subunit and a large subunit. Both
 CC participate in formation of the active center, but the ATP-binding
 CC site is exclusively located on the small subunit (By similarity).
 CC CC -!- SIMILARITY: Belongs to the eukaryotic-type primase large subunit
 CC family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AJ248283; CAB49107.1; -;
 CC PIR; D75207; D75207.
 CC HAVAP; MF_00701; -; 1.
 CC InterPro; IPR007238; DNA primase lrg.
 CC Pfam; PF04104; DNA primase lrg; 1.
 CC Transferrase; DNA replication; DNA-directed RNA polymerase; Primosome;
 CC

KW Complete proteome.
SQ SEQUENCE 393 AA; 45485 MW; 9448642PBB3PBBF43 CRC64;
Query Match 75.5%; Score 37; DB 1; Length 393;
Best Local Similarity 66.7%; Pred.No. 5.2;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 KMEYTVHL 9
DB 129 KLPYTIHL 137

RESULT 6
MID1 HUMAN
ID MID1 HUMAN STANDARD; PRT; 667 AA.
AC O15344; O75361; Q9BZX5;
DT 28-FEB-2003 (Rel. 41; Created)
DT 28-FEB-2003 (Rel. 41; Last sequence update)
DT 15-MAR-2004 (Rel. 43; Last annotation update)
DE Midline 1 protein (Tripartite motif protein 18) (Putative
DE transcription factor XPRF).
OS MID1 OR FX1 OR TRIM18 OR XPRF.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANTS OS-I MET-458 DEL AND
RP PHE-ILE-ASP-SER-GLY-ARG-HIS-LEU-534 INS.
RC TISSUE=Neuron;
RX MEDLINE=98016411; PubMed=9354791;
RA Quaderi N., Schweiger S., Gaudenz K., Franco B., Rugarli E.I.,
RA Berger W., Feldman C.J., Volta M., Andolfi G., Gilgenkrantz S.,
RA Marion R.W., Hennekam R.C.M., Opitz J.M., Muenke M., Ropers H.H.,
RA Ballabio A.;
RT "Optiz G/BBB syndrome, a defect of midline development, is due to
RT mutations in a new RING finger gene on Xp22.";
RL Nat. Genet. 17:285-291(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Placenta;
RX MEDLINE=98087583; PubMed=9425238;
RA Perry J., Feather S., Smith A., Palmer S., Ashworth A.;
RT "The human FX1 gene is located within Xp22.3: implications for
RT evolution of the mammalian X chromosome.";
RL Hum. Mol. Genet. 7:299-305(1998).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Fetal kidney;
RX MEDLINE=98390188; PubMed=9722948;
RA Van den Veyver I.B., Cormier T.A., Jurecic V., Baldini A.,
RA Zoghbi H.Y.;
RT "Characterization and physical mapping in human and mouse of a novel
RT RING finger gene in Xp22.";
RL Genomics 51:251-261(1998).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT OS-I PRO-626.
RC TISSUE=Brain;
RX MEDLINE=20485428; PubMed=11030761;
RA Cox T.C., Allen L.R., Cox L.L., Hopwood B., Goodwin B., Haan E.,
RA Suthers G.K.;
RT "New mutations in MID1 provide support for loss of function as the
RT cause of X-linked opitz syndrome.";
RL Hum. Mol. Genet. 9:2553-2562(2000).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX MEDLINE=21231161; PubMed=11331580;
RA Raymond A., Meroni G., Fantozzi A., Merla G., Cairo S., Luzzi L.,
RA Riganelli D., Zanaria E., Messali S., Gainarca S., Guffanti A.,
RA Minucci S., Pelicci P.G., Ballabio A.;
RT "The tripartite motif family identifies cell compartments.";
RL EMBO J. 20:2140-2151(2001).
RN [6]

RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Berge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [7]
RP INTERACTION WITH IGBP1, AND PHOSPHORYLATION.
RX PubMed=11806752;
RA Short K.M., Hopwood B., Yi Z., Cox T.C.;
RT "MID1 and MID2 homo- and heterodimerise to tether the
RT rapamycin-sensitive PP2A regulatory subunit, Alpha 4, to
RT microtubules: implications for the clinical variability of X-linked
RT Opitz GBBB syndrome and other developmental disorders.";
RL BMC Cell Biol. 3:1-1(2002).
RN [8]
RP VARIANTS OS-I ARG-266 AND THR-536.
RX MEDLINE=98386064; PubMed=9718340;
RA Gaudenz K., Quaderi N., Franco B., Feldman G.,
RA Gasser D.L., Wittwer B., Horst J., Montini E., Opitz J.M.,
RA Ballabio A., Muenke M.;
RA "Opitz G/BBB syndrome in Xp22: mutations in the MID1 gene cluster in
RT the carboxy-terminal domain.";
RL Am. J. Hum. Genet. 63:703-710(1998).
RN [9]
RP ERRATUM.
RA Gaudenz K., Roessler E., Quaderi N., Franco B., Feldman G.,
RA Gasser D.L., Wittwer B., Horst J., Montini E., Opitz J.M.,
RA Ballabio A., Muenke M.;
RL Am. J. Hum. Genet. 63:1571-1571(1998).
CC -!- FUNCTION: Putative transcriptional regulator.
CC -!- SUBUNIT: Homodimer or heterodimer with MID2. Interacts with IGBP1.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=Alpha;
CC IsoId=O15344-1; Sequence=Displayed;
CC Name=2; Synonyms=Beta;
CC IsoId=O15344-2; Sequence=VSP_005735;
CC -!- TISSUE SPECIFICITY: In the fetus, highest expression found in
CC kidney, followed by brain and lung. Expressed at low levels in
CC fetal liver. In the adult, most abundant in heart, placenta and
CC brain.
CC -!- DISEASE: Defects in MID1 are the cause of Opitz syndrome type I
CC (OS-I) [MIM:300000]. OS-I is an X-linked recessive disorder
CC characterized by hypotelorism, genital-urinary defects such as
CC hypospadias in males and splayed labia in females, lip-palate-
CC laryngotracheal clefts, imperforate anus, developmental delay and
CC congenital heart defects.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC -!- SIMILARITY: Contains 1 B box-type zinc finger.
CC -!- SIMILARITY: Contains 1 fibronectin type III domain.
CC -!- SIMILARITY: Contains 1 SPR domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL; Y13667; CAA74018.1; -
CC EMBL; AF035360; AAB99951.1; -
CC EMBL; AF041206; AAC32998.1; -
CC EMBL; AF041207; AAC32999.1; -
CC EMBL; AF041208; AAC33000.1; -
CC EMBL; AF041209; AAC33001.1; -
CC EMBL; AF041210; AAC33002.1; -
CC EMBL; AF230976; AAG50191.1; -
CC EMBL; AF230977; AAG50192.1; -
CC EMBL; AF269101; AAG33130.1; -
CC EMBL; BC053626; AAH53626.1; -
CC PIR; T09482; T09482.
CC Genew; HGNC:7095; MID1.
CC MIM; 300000; -
CC DR GO; GO:0005875; C:Microtubule associated complex; TAS.
CC DR GO; GO:0005515; F:Protein binding; TAS.
CC DR GO; GO:0000226; P:Microtubule cytoskeleton organization and b. .; TAS.
CC DR GO; GO:0007389; P:Pattern specification; TAS.
CC DR InterPro; IPR001870; B302.
CC DR InterPro; IPR003649; Bbox.C.
CC DR InterPro; IPR008957; FN.III-like.
CC DR InterPro; IPR003961; FN.III.
CC DR InterPro; IPR006574; PRY.
CC DR InterPro; IPR003877; SPRY receptor.
CC DR InterPro; IPR000315; Znf.Bbox.
CC DR InterPro; IPR001841; Znf.Ring.
CC DR Pfam; PF00041; FN3.1.
CC DR Pfam; PF00622; SPRY.1.
CC DR Pfam; PF00643; Zf-B.Box.2.
CC DR Pfam; PF00097; Zf-C3HC4.1.
CC DR SMART; SM00502; BBC.1.
CC DR SMART; SM00336; BBOX.2.
CC DR SMART; SM00060; FN3.1.
CC DR SMART; SM00589; PRY.1.
CC DR SMART; SM00184; RING.1.
CC DR SMART; SM00449; SPRY.1.
CC DR PROSITE; PS00119; ZF.BBOX.1.
CC DR PROSITE; PS00518; ZF.RING.1.
CC DR PROSITE; PS00089; ZF.RING.2.
CC DR Transcription regulation; Nuclear protein; Coiled coil; Zinc-finger;
CC Metal-binding.
CC ZN.FING 10 60
CC ZN.FING 170 212
CC DOMAIN 205 264
CC DOMAIN 384 477
CC DOMAIN 538 656
CC VARSPLIC 553 667
CC Missing (in isoform 2).
CC /FTid=VSP_005735.
CC C -> R (in OS-I).
CC /FTid=VAR_013758.
CC Missing (in OS-I).
CC /FTid=VAR_013759.
CC V -> VFDISGRHL (in OS-I).
CC /FTid=VAR_013760.
CC I -> T (in OS-I).
CC /FTid=VAR_013761.
CC L -> P (in OS-I).
CC /FTid=VAR_013762.
CC T -> P (IN REF. 3; AAC32999).
CC CONFLICT 228 228
CC CONFLICT 484 484
CC Q -> P (IN REF. 3; AAC32998).
CC SEQUENCE 667 AA; 75250 MW; 673C5120018BA619 CRC64;

Query Match 71.4%; Score 35; DB 1; Length 667;
Best Local Similarity 75.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 K0NEYTVH 8

Db 444 K0NEYTVH 451
RESULT 7
MIDI_MUASP
ID MIDI_MUASP STANDARD; PRT; 667 AA.
AC P82457;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Midline 1 protein (Tripartite motif protein 18).
DE MIDI OR FXY OR TRIM18.
OS Mus spretus (Western wild mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10096;
RN [1]
RP SEQUENCE FROM N.A. PubMed=10508587;
RX MEDLINE=99439873; PubMed=10508587;
RA Perry J., Ashworth A.;
RT "Evolutionary rate of a gene affected by chromosomal position.";
RL Curr. Biol. 9:987-989(1999).
CC -!- FUNCTION: Putative transcriptional regulator.
CC -!- SUBUNIT: Homodimer or heterodimer with MID2. Interacts with IGBP1
CC (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC -!- SIMILARITY: Contains 1 B box-type zinc finger.
CC -!- SIMILARITY: Contains 1 fibronectin type III domain.
CC -!- SIMILARITY: Contains 1 SPRY domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL; AF186460; AAD56246.1; -
CC MGD; MGI:1100537; Mid1.
CC InterPro; IPR001870; B302.
CC InterPro; IPR003649; Bbox.C.
CC InterPro; IPR008957; FN.III-like.
CC InterPro; IPR003961; FN.III.
CC InterPro; IPR006574; PRY.
CC InterPro; IPR003877; SPRY receptor.
CC InterPro; IPR000315; Znf.Bbox.
CC InterPro; IPR001841; Znf.Ring.
CC Pfam; PF00041; FN3.1.
CC Pfam; PF00643; Zf-B.Box.2.
CC Pfam; PF00097; Zf-C3HC4.1.
CC SMART; SM00502; BBC.1.
CC SMART; SM00336; BBOX.2.
CC SMART; SM00060; FN3.1.
CC SMART; SM00589; PRY.1.
CC SMART; SM00184; RING.1.
CC SMART; SM00449; SPRY.1.
CC PROSITE; PS00119; ZF.BBOX.1.
CC PROSITE; PS00518; ZF.RING.1.
CC PROSITE; PS00089; ZF.RING.2.
CC Transcription regulation; Nuclear protein; Coiled coil; Zinc-finger;
CC Metal-binding.
CC ZN.FING 10 60
CC ZN.FING 170 212
CC DOMAIN 205 264
CC DOMAIN 384 477
CC DOMAIN 538 656
CC SEQUENCE 667 AA; 75173 MW; D6EFA1CDEA43CB9 CRC64;
Query Match 71.4%; Score 35; DB 1; Length 667;

Best Local Similarity 75.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KNEYTYVH 8
Db 444 KQHYTVH 451

RESULT 8
MID1_RAT
ID MID1_RAT STANDARD; PRT; 667 AA.
AC P82456;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Midline 1 protein (Tripartite motif protein 18).
GN MID1 OR FXY OR TRIM18.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99439873; PubMed=10508587;
RA Perry J., Ashworth A.;
RT "Evolutionary rate of a gene affected by chromosomal position.";
RL Curr. Biol. 9:987-989(1999).
CC -!- FUNCTION: Putative transcriptional regulator.
CC -!- SUBUNIT: Homodimer or heterodimer with MID2. Interacts with IGBP1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC -!- SIMILARITY: Contains 1 B box-type zinc finger.
CC -!- SIMILARITY: Contains 1 fibronectin type III domain.
CC -!- SIMILARITY: Contains 1 SPRY domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; AF186461; AAD56247.1; -;
CC InterPro; IPR001870; B302.
CC InterPro; IPR003649; BBox_C.
CC InterPro; IPR008957; FN III-like.
CC InterPro; IPR003961; FN III.
CC InterPro; IPR008574; PRY.
CC InterPro; IPR003877; SPRY_receptor.
CC InterPro; IPR000315; Znf_Bbox.
CC Pfam; PF00041; fn3; 1.
CC Pfam; PF00622; SPRY; 1.
CC Pfam; PF00643; zf-B_box; 2.
CC Pfam; PF00097; zf-C3HC4; 1.
CC SMART; SM00502; BBox; 1.
CC SMART; SM00336; BBox; 2.
CC SMART; SM00060; FN3; 1.
CC SMART; SM00589; PRY; 1.
CC SMART; SM00184; RING; 1.
CC SMART; SM00449; SPRY; 1.
CC PROSITE; PS50119; 2F_BBOX; 1.
CC PROSITE; PS00518; 2F_RING_1; 1.
CC PROSITE; PS50089; 2F_RING_2; 1.
KW Transcription regulation; Nuclear protein; Coiled coil; Zinc-finger; Metal-binding.
FT ZN_FING 10 60 RING-TYPE.
FT ZN_FING 170 212 B BOX-TYPE.
FT DOMAIN 205 264 COILED COIL (POTENTIAL).
FT DOMAIN 384 477 FIBRONECTIN TYPE-III.
FT DOMAIN 538 657 SPRY.

SQ SEQUENCE 667 AA; 75210 MW; BA73528FEAE59603 CRC64;
Query Match 71.4%; Score 35; DB 1; Length 667;
Best Local Similarity 75.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KNEYTYVH 8
Db 444 KQHYTVH 451

RESULT 9
MID1_MOUSE
ID MID1_MOUSE STANDARD; PRT; 680 AA.
AC Q70583; O35418;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Midline 1 protein (Tripartite motif protein 18).
GN MID1 OR FXY OR TRIM18.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX TISSUE=Embryo;
RX MEDLINE=98133931; PubMed=9467009;
RA Zotto L.D., Quaderi N.A., Elliott R., Lingerfelter P.A., Carrel L., Valasecchi V., Monti B., Yen C.-H., Chapman V., Kaicheva I., Arrigo G., Zuffardi O., Thomas S., Willard H.F., Ballabio A., Distèche C.M., Rugardi E.I.;
RT "The mouse Mid1 gene: implications for the pathogenesis of Opitz syndrome and the evolution of the mammalian pseudoautosomal region.";
RL Hum. Mol. Genet. 7:489-499(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=98004518; PubMed=9342357;
RA Palmer S., Perry J., Kipling D., Ashworth A.;
RT "A gene spans the pseudoautosomal boundary in mice.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:12030-12035(1997).
RN [3]
RP TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RX MEDLINE=98390188; PubMed=9722948;
RA Van den Veyver I.B., Cormier T.A., Jurecic V., Baldini A., Zoghbi H.Y.;
RT "Characterization and physical mapping in human and mouse of a novel RING finger gene in Xp22.";
RL Genomics 51:251-261(1998).
CC -!- FUNCTION: Putative transcriptional regulator.
CC -!- SUBUNIT: Homodimer or heterodimer with MID2. Interacts with IGBP1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O70583-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O70583-2; Sequence=VSP_005736;
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed in fetus and adult. At E9-E10.5, highest expression found in frontonasal processes, branchial arches and CNS. From E12.5 to E16.5, high levels found in rostral part of CNS. At E14.5, begins to be highly expressed in kidney and lung. At E16.5, highly expressed in the mucosa of the hindgut and cutaneous region of the stomach.
CC -!- DEVELOPMENTAL STAGE: Expressed throughout embryonic development with highest levels from E7-E11. Also expressed in the adult.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC -!- SIMILARITY: Contains 1 B box-type zinc finger.
CC -!- SIMILARITY: Contains 1 fibronectin type III domain.
CC -!- SIMILARITY: Contains 1 SPRY domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; Y14848; CAA75113.1; -;
 CC DR EMBL; AF026565; AAB83986.1; -;
 CC DR MGD; MGI:1100537; Mid1.
 CC DR InterPro; IPR001870; B302.
 CC DR InterPro; IPR003649; Bbox C.
 CC DR InterPro; IPR008957; FN-III-like.
 CC DR InterPro; IPR003961; FN-III.
 CC DR InterPro; IPR003877; SPRY receptor.
 CC DR InterPro; IPR000315; Znf Bbox.
 CC DR InterPro; IPR001841; Znf_ring.
 CC DR Pfam; PF00041; fn3; 1.
 CC DR Pfam; PF00622; SPRY; 1.
 CC DR Pfam; PF00643; zf-B-box; 2.
 CC DR Pfam; PF00097; zf-C3HC4; 1.
 CC DR SMART; SM00502; BBC; 1.
 CC DR SMART; SM00336; BBOX; 2.
 CC DR SMART; SM00060; FN3; 1.
 CC DR SMART; SM00184; RING; 1.
 CC DR SMART; SM00449; SPRY; 1.
 CC DR PROSITE; PS00119; ZF-BBOX; 1.
 CC DR PROSITE; PS00518; ZF_RING_1; 1.
 CC DR PROSITE; PS00089; ZF_RING_2; 1.
 CC DR Transcription regulation; Nuclear protein; Coiled coil; Zinc-finger;
 CC Metal-binding; Alternative splicing.
 CC KW ZN_FING 10 60
 CC FT ZN_FING 170 212
 CC FT B_BOX-TYPE
 CC FT DOMAIN 205 264
 CC FT COILED COIL (POTENTIAL).
 CC FT DOMAIN 384 490
 CC FT FIBRONECTIN TYPE-III.
 CC FT DOMAIN 551 669
 CC FT VARSPIC 429 442
 CC FT NVACDGTCLGSAG -> S (in isoform 2).
 CC FT /FTID-VSP 005736.
 CC FT M -> T (IN REF. 2).
 CC FT CONFLICT 174 174
 CC FT CONFLICT 523 523
 CC FT CONFLICT 633 633
 CC FT A -> B (IN REF. 2).
 CC FT A -> T (IN REF. 2).
 CC FT SEQUENCE 680 AA; 76121 MW; A9C2E5149A0F7802 CRC64;
 CC

Query Match 71.4%; Score 35; DB 1; Length 680;
 Best Local Similarity 75.0%; Pred. No. 24;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KQNEYTVH 8
 DB 457 KQNHVTVH 464

RESULT 10
 MID2 MOUSE
 ID MID2 MOUSE STANDARD; PRT; 685 AA.
 AC OSQUS6;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Midline 2 protein (Midline defect 2) (Tripartite motif protein 1).
 GN MID2 OR FX2 OR TRIM1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 OC [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99330546; PubMed=10400986;
 RA Buchner G., Montini E., Andolfi G., Quaderi N., Cainarca S.,
 RA Messali S., Bassi M.T., Ballabio A., Meroni G., Franco B.,
 RT "MID2, a homologue of the Optiz syndrome gene MID1: similarities in a
 RT sub-cellular localization and differences in expression during
 RT development.";

RL Hum. Mol. Genet. 8:1397-1407(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20112752; PubMed=10644436;
 RA Perry J., Short K.M., Romer J.T., Swift S., Cox T.C., Ashworth A.;
 RT "FX2/MID2, a gene related to the X-linked Opitz syndrome gene
 RT FX2/MID1, maps to Xq22 and encodes a FNIII domain-containing protein
 RT that associates with microtubules.";
 RL Genomics 62:385-394(1999).
 CC -!- SUBUNIT: Homodimer or heterodimer with MID1. Interacts with IGBP1
 CC (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic; microtubule-associated.
 CC -!- TISSUE SPECIFICITY: Low abundance in brain and lung, with even
 CC lower levels in heart, liver, and kidney.
 CC -!- DEVELOPMENTAL STAGE: At E10.5, a very low level is mostly confined
 CC to the central nervous system and the developing heart and kidney,
 CC while at later stages it is present in other organ systems.
 CC -!- DOMAIN: The tripartite motif (RBC; RING- and B-box-type zinc
 CC fingers and coiled coil domains) mediates dimerization (By
 CC similarity).
 CC -!- DOMAIN: Associates with microtubules in a manner that is dependent
 CC on the C-terminal B30.2 domain.
 CC -!- PTM: Phosphorylated on serine and threonine residues (By
 CC similarity).
 CC -!- SIMILARITY: Belongs to the TRIM/RBCC family.
 CC -!- SIMILARITY: Contains 1 B30.2-like domain.
 CC -!- SIMILARITY: Contains 2 B-box-type zinc fingers.
 CC -!- SIMILARITY: Contains 1 fibronectin type III domain.
 CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; Y18881; CAB56170.1; -;
 EMBL; AF196480; AAF07340.1; -;
 MGD; MGI:1344333; Mid2.
 InterPro; IPR001870; B302.
 InterPro; IPR003649; Bbox C.
 InterPro; IPR003961; FN-III.
 InterPro; IPR008957; FN-III-like.
 InterPro; IPR003877; SPRY receptor.
 InterPro; IPR000315; Znf Bbox.
 InterPro; IPR001841; Znf_ring.
 Pfam; PF00041; fn3; 1.
 Pfam; PF00622; SPRY; 1.
 Pfam; PF00643; zf-B-box; 1.
 Pfam; PF00097; zf-C3HC4; 1.
 SMART; SM00502; BBC; 1.
 SMART; SM00336; BBOX; 2.
 SMART; SM00060; FN3; 1.
 SMART; SM00184; RING; 1.
 SMART; SM00449; SPRY; 1.
 PROSITE; PS00119; ZF-BBOX; 1.
 PROSITE; PS00518; ZF_RING_1; 1.
 PROSITE; PS00089; ZF_RING_2; 1.
 KW Microtubule; Repeat; Coiled coil; Metal-binding; Zinc-finger; Zinc;
 KW Phosphorylation.
 FT ZN_FING 10 60
 FT ZN_FING 114 164
 FT ZN_FING 170 212
 FT B_BOX-TYPE 1.
 FT ZN_FING 170 212
 FT COILED COIL (POTENTIAL).
 FT DOMAIN 208 275
 FT FIBRONECTIN TYPE-III.
 FT DOMAIN 384 477
 FT DOMAIN 486 641
 FT B30.2-LIKE.
 SQ SEQUENCE 685 AA; 77841 MW; ODA2386C004909A0 CRC64;
 Query Match 71.4%; Score 35; DB 1; Length 685;
 Best Local Similarity 75.0%; Pred. No. 24;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 K0NEYTVH 8
Db 444 K0NHYTVH 451

RESULT 11
MID2 HUMAN
ID MID2 HUMAN STANDARD; PRT; 715 AA.
AC Q9UJV3; Q8WVK1; Q9UUR9;
CT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Midline 2 protein (Midline defect 2) (tripartite motif protein 1)
DE (Midin 2).
GN MID2 OR FXY2 OR TRIM1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=20112752; PubMed=10644436;
RA Perry J., Short K.M., Romer J.T., Swift S., Cox T.C., Ashworth A.,
RT "FX2/MID2, a gene related to the X-linked Opitz syndrome gene
RT FX2/MID1, maps to Xq22 and encodes a FNIII domain-containing protein
RT that associates with microtubules.";
RL Genomics 62:395-394(1999).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX TISSUE=Fetal brain;
RA Buchner G., Montini E., Andolfi G., Quaderi N., Cainarca S.,
RA Messali S., Bassi M.T., Ballabio A., Meroni G., Franco B.,
RT "MID2, a homologue of the Opitz syndrome gene MID1: similarities in a
RT sub-cellular localization and differences in expression during
RT development.";
RL Hum. Mol. Genet. 8:1397-1407(1999).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RA Kalnine N., Chen X., Rolf A., Halleck A., Hines L., Eisenstein S.,
RA Koundinya M., Raphael J., Moreira D., Kelley T., Labaer J., Lin Y.,
RA Phelan M., Farmer A.,
RT "Cloning of human full-length cDNAs in BD Creator(TM) system donor
RT vector.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX TISSUE=Kidney;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Iqbal N.A., Peters G.J., Abramson R.D., Mullen S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
RN [5]
RP SEQUENCE OF 1-252 FROM N.A.
RA Heath P.;

Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
[6]
RN INTERACTION WITH IGBP1, AND PHOSPHORYLATION.
RP PubMed=11806752;
RX Short K.M., Hopwood B., Yi Z., Cox T.C.;
RT "MID1 and MID2 homo- and heterodimerise to tether the
RT rapamycin-sensitive PP2A regulatory subunit, Alpha 4, to
RT microtubules: implications for the clinical variability of X-linked
RT Opitz GBBB syndrome and other developmental disorders.";
RL BMC Cell Biol. 3:1-1(2002)
CC [1]
CC SUBUNIT: Homodimer or heterodimer with MID1. Interacts with IGBP1.
CC [2]
CC SUBCELLULAR LOCATION: Cytoplasmic; microtubule-associated.
CC [3]
CC ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9UJV3-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9UJV3-2; Sequence=VSP_009009;
CC [4]
CC TISSUE SPECIFICITY: Low level in fetal kidney and lung, and in
CC adult prostate, ovary and small intestine.
CC [5]
CC DOMAIN: The tripartite motif (RBCC; RING- and B box-type zinc
CC fingers and coiled coil domains) mediates dimerization (By
CC similarity).
CC [6]
CC DOMAIN: Associates with microtubules in a manner that is dependent
CC on the C-terminal B30.2 domain.
CC [7]
CC PTM: Phosphorylated on serine and threonine residues.
CC [8]
CC SIMILARITY: Belongs to the TRIM/RBCC family.
CC [9]
CC SIMILARITY: Contains 1 B30.2-like domain.
CC [10]
CC SIMILARITY: Contains 2 B box-type zinc fingers.
CC [11]
CC SIMILARITY: Contains 1 fibronectin type III domain.
CC [12]
CC SIMILARITY: Contains 1 RING-type zinc finger.
CC [13]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch)
CC [14]
CC EMBL; AF196481; AAF07341.1; -
CC EMBL; Y18880; CABS6154.1; -
CC EMBL; BT006663; AAP35309.1; -
CC EMBL; BC017707; AAHI7707.1; -
CC EMBL; AL109946; CAD19102.1; -
CC Genew; HGNC:7096; MID2.
CC MIM; 300204; -
CC InterPro; IPR001870; B302.
CC InterPro; IPR003849; BBox_C.
CC InterPro; IPR003961; FN_III.
CC InterPro; IPR008957; FN_III-like.
CC InterPro; IPR003877; SPRY receptor.
CC InterPro; IPR000315; Znf BBox.
CC InterPro; IPR001841; Znf_ring.
CC Pfam; PF00041; fn3; 1.
CC Pfam; PF06622; SPRY; 1.
CC Pfam; PF06643; ZF-B_Box; 1.
CC Pfam; PF00097; ZF-C3HC4; 1.
CC PRINTS; PR01406; BBOXZNFINGER.
CC SMART; SM00502; BBC; 1.
CC SMART; SM00336; BBOX; 2.
CC SMART; SM00060; FN3; 1.
CC SMART; SM00184; RING; 1.
CC SMART; SM00449; SPRI; 1.
CC PROSITE; PS00119; ZF_BBOX; 1.
CC PROSITE; PS00518; ZF_RING_1; 1.
CC PROSITE; PS00089; ZF_RING_2; 1.
CC Microtubule; Repeat; Coiled coil; Metal-binding; Zinc-finger; Zinc;
CC Phosphorylation; Alternative splicing.
FT ZN_FING 10 60 RING-TYPE
FT ZN_FING 114 164 B BOX-TYPE 1.
FT ZN_FING 170 212 B BOX-TYPE 2.
FT DOMAIN 208 275 COILED COIL (POTENTIAL).

```

FT DOMAIN 384 507 FIBRONECTIN TYPE-III.
FT DOMAIN 516 671 B30.2-LIKE.
FT VARSPLIT 430 459 Missing (in isoform 2).
FT /FTid=VSP_009009.
SQ SEQUENCE 715 AA; 81263 MW; B10518806E358BFC CRC64;

Query Match 71.4%; Score 35; DB 1; Length 715;
Best Local Similarity 75.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KNNHYTVH 8
| | | | |
Db 474 KNNHYTVH 481

RESULT 12
RFA2_RAT STANDARD; PRT; 266 AA.
AC Q63528;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Replication protein A 32 kDa subunit (RP-A) (RP-A) (Replication
DE factor-A protein 2) (Fragment).
GN RPA2.
OS Rattus norvegicus [Rat].
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Nagelhus T., Haug T., Krokan H.E.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Absolutely required for simian virus 40 DNA replication
CC in vitro. It participates in a very early step in initiation. RP-A
CC is a single-stranded DNA-binding protein.
CC -!- SUBUNIT: Heterotrimer of 70, 32, and 14 kDa chains. The DNA-
CC binding activity may reside exclusively on the 70 kDa subunit.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- PTM: Phosphorylated in a cell-cycle-dependent manner (from the S
CC phase until mitosis) (By similarity).
CC -!- SIMILARITY: TO OTHER SPECIES RFA2/RPA2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X98490; CAA67116.1; -.
CC InterPro; IPR008994; Nucleic_acid_OB.
CC InterPro; IPR004365; tRNA anti.
CC Pfam; PF01336; tRNA anti. 1.
CC DNA replication; Phosphorylation; Nuclear protein.
FT NON TPR 1
SQ SEQUENCE 266 AA; 28827 MW; 52E13704C1D2E13A CRC64;

Query Match 69.4%; Score 34; DB 1; Length 266;
Best Local Similarity 62.5%; Pred. No. 14;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 MNEYTVHL 9
| | | | |
Db 148 MNEYTAHI 155

RESULT 13
RFA2_HUMAN STANDARD; PRT; 270 AA.
AC P15927;
DT 01-APR-1990 (Rel. 14, Created)

```

```

DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Replication protein A 32 kDa subunit (RP-A) (RP-A) (Replication
DE factor-A protein 2).
GN RPA2 OR RPA32 OR RPA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX MEDLINE=90153966; PubMed=2406247;
RA Erdlie L.F., Wold M.S., Kelly T.J.;
RT "The primary structure of the 32-kDa subunit of human replication
RT protein A.";
RL J. Biol. Chem. 265:3177-3182(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney, Lung, and Muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Souffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Absolutely required for simian virus 40 DNA replication
CC in vitro. It participates in a very early step in initiation. RP-A
CC is a single-stranded DNA-binding protein.
CC -!- SUBUNIT: Heterotrimer of 70, 32/30, and 14 kDa chains. The DNA-
CC binding activity may reside exclusively on the 70 kDa subunit.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- PTM: Phosphorylated in a cell-cycle-dependent manner (from the S
CC phase until mitosis).
CC -!- SIMILARITY: TO OTHER SPECIES RFA2/RPA2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J05249; AAA36560.1; -.
CC EMBL; BC012157; AAH12157.1; -.
CC EMBL; BC001630; AAH01630.1; -.
CC EMBL; BC021257; AAH21257.1; -.
CC PIR; A43711; A43711.
CC FDB; IDPU; 10-NOV-00.
CC FDB; ILIO; 05-JUN-02.
CC FDB; IQUC; 02-SEP-99.
CC Genew; HGNC:10290; RPA2.
CC GK; P15927; -.
CC MIM; 179836; -.
CC GO; GO:0005662; C:DNA replication factor A complex; TAS.
CC GO; GO:0003697; F:single-stranded DNA binding; TAS.
CC GO; GO:0006261; F:DNA dependent DNA replication; TAS.
CC InterPro; IPR008994; Nucleic_acid_OB.

```

DR InterPro; IPR004365; tRNA anti.
DR Pfam; PF01336; tRNA anti.
KW DNA replication; Phosphorylation; Nuclear protein; 3D-structure.
FT DOMAIN 1 29
FT DOMAIN 37 45
FT DOMAIN 95 123
FT DOMAIN 127 145
FT DOMAIN 247 270
SQ SEQUENCE 270 AA; 29247 MW; 61A563EA7B34A9B1 CRC64;

Query Match 69.4%; Score 34; DB 1; Length 270;
Best Local Similarity 62.5%; Pred. No. 14;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 MNEVTVHL 9
Db 152 MNEFTTHI 159

RESULT 14
RFA2_MOUSE STANDARD; PRT; 270 AA.
AC Q62193;
DT 01-NOV-1997 (Rel. 35; Created)
DT 01-NOV-1997 (Rel. 35; Last sequence update)
DT 30-MAY-2000 (Rel. 39; Last annotation update)
DE Replication protein A 32 kDa subunit (RP-A) (RP-A) (Replication factor-A protein 2).
GN RPA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91334146; PubMed=1908076;
RA Nakagawa M., Tsukada S., Soma T., Shimizu Y.K., Miyake S.,
RA Iwanatsu A., Sugiyama H.;
RT "cDNA cloning of the murine 30-kDa protein homologous to the 32-kDa subunit of human replication protein A";
RL Nucleic Acids Res. 19:4292-4292(1991).
CC -!- FUNCTION: Absolutely required for simian virus 40 DNA replication in vitro. It participates in a very early step in initiation. RP-A is a single-stranded DNA-binding protein.
CC -!- SUBUNIT: Heterotrimer of 70, 32, and 14 kDa chains. The DNA-binding activity may reside exclusively on the 70 kDa subunit.
CC -!- PTM: Phosphorylated in a cell-cycle-dependent manner (from the S phase until mitosis) (by similarity).
CC -!- SIMILARITY: TO OTHER SPECIES RFA2/RPA2.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D00812; BAA00693.1; -
DR PIR; S28682; S28682.
DR MGD; MGI:133939; Rpa2.
DR InterPro; IPR008994; Nucleic acid OB.
DR InterPro; IPR004365; tRNA anti.
DR Pfam; PF01336; tRNA anti;
KW DNA replication; Phosphorylation; Nuclear protein.
SQ SEQUENCE 270 AA; 29718 MW; BF0EF86612A48011 CRC64;

Query Match 69.4%; Score 34; DB 1; Length 270;
Best Local Similarity 62.5%; Pred. No. 14;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 MNEVTVHL 9

Db 152 MNEFTTHI 159

RESULT 15
Y539_AQUAE STANDARD; PRT; 285 AA.
ID Y539_AQUAE
AC O66818;
DT 15-OCT-2001 (Rel. 40; Created)
DT 15-OCT-2001 (Rel. 40; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Hypothetical protein AQ_539.
GN AQ_539.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Shead M.A., Keller M., Rujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex aeolicus";
RL Nature 392:353-358(1998).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AE000694; AAC06778.1; -
DR PIR; F70348; F70348.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 7 29
FT TRANSMEM 49 71
FT TRANSMEM 95 117
FT TRANSMEM 137 156
FT TRANSMEM 232 254
FT TRANSMEM 259 281
SQ SEQUENCE 285 AA; 33543 MW; FDC54632F0EF39BB CRC64;

Query Match 69.4%; Score 34; DB 1; Length 285;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KMNEVTVHL 9
Db 207 KNEVTLNL 215

Search completed: March 1, 2004, 17:29:53
Job time : 8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:20:41 ; Search time 30.3333 Seconds

(without alignments)
93.615 Million cell updates/sec

Title: US-09-905-083-31

Perfect score: 49

Sequence: 1 KXNEYTVHL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_virus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	100.0	253	4 Q8N5N9	Q8N5N9 homo sapien
2	44	89.8	181	4 Q8NFV7	Q8NFV7 homo sapien
3	39	79.6	210	9 Q8G0B3	Q8G0B3 bacterioph
4	37	75.5	334	16 Q8D5G2	Q8D5G2 streptococ
5	36	73.5	603	10 Q8L836	Q8L836 arabidopsis
6	36	73.5	627	5 Q9XU04	Q9XU04 caenorhabd
7	36	73.5	635	10 Q9LQW1	Q9LQW1 arabidopsis
8	36	73.5	753	16 Q8EGM8	Q8EGM8 streptococ
9	36	73.5	753	16 Q8E174	Q8E174 streptococ
10	36	73.5	753	16 Q8VD00	Q8VD00 streptococ
11	36	73.5	1829	5 Q19815	Q19815 caenorhabd
12	35	71.4	47	11 Q810Y2	Q810Y2 mus spicile
13	35	71.4	47	11 Q80ZB4	Q80ZB4 mus platyth
14	35	71.4	47	11 Q80Z83	Q80Z83 arvicola te
15	35	71.4	47	11 Q80Z80	Q80Z80 mastomys nu
16	35	71.4	47	11 Q80Z77	Q80Z77 mus cervico

17 35 71.4 47 11 Q80Z74
18 35 71.4 48 11 Q810X9
19 35 71.4 159 16 Q82711
20 35 71.4 159 16 Q8Y419
21 35 71.4 197 12 Q875S2
22 35 71.4 505 16 Q8RV79
23 35 71.4 555 16 Q826Z1
24 35 71.4 579 5 Q81BM5
25 35 71.4 667 13 Q80WD1
26 35 71.4 685 4 Q8UJV3
27 35 71.4 687 5 Q81E27
28 35 71.4 715 4 Q8UJR9
29 35 71.4 1408 5 Q81D31
30 35 71.4 2213 5 Q81HW8
31 35 71.4 270 11 Q89KL9
32 34 69.4 299 17 Q8TPI9
33 34 69.4 421 16 Q82ES2
34 34 69.4 421 16 Q8YAO0
35 34 69.4 1179 5 Q81DD1
36 34 69.4 1946 5 Q87291
37 34 67.3 182 16 Q81AA5
38 33 67.3 211 16 Q7VGR8
39 33 67.3 213 16 Q8F4B8
40 33 67.3 218 16 Q8K888
41 33 67.3 227 16 Q9KON9
42 33 67.3 281 16 Q892A6
43 33 67.3 320 16 Q814B3
44 33 67.3 327 3 Q86V06
45

ALIGNMENTS

RESULT 1

Q8N5N9 PRELIMINARY; PRT; 253 AA.
AC Q8N5N9
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Kallikrein 7 (chymotryptic, stratum corneum).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; BC032005; AAH32005.1; -
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser_trypsin.
DR InterPro; IPR001354; Peptidase_S1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 253 AA; 27608 MW; 2D68B6A1B22A668 CRC64;

Query Match 100.0%; Score 49; DB 4; Length 253;

Best Local Similarity 100.0%; Pred.No. 0.076;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KNEYTVHL 9
Db 72 KNEYTVHL 80

RESULT 2
Q8NFV7 PRELIMINARY; PRT; 181 AA.
AC Q8NFV7;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Kallikrein 7 short variant protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Ovarian Carcinoma;
RA Dong Y., Kaushal A., Clements J.A.;
RT "Human Kallikrein 7 (KLK7) short variant mRNA from ovarian
carcinoma";
RT Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RL -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC EMBL; AF411215; AAN03663.1; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004298; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRY-SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydroxylase; Protease; Serine protease.
SQ SEQUENCE 181 AA; 19887 MW; 86A28A03B80C2D78 CRC64;

Query Match 89.8%; Score 44; DB 4; Length 181;
Best Local Similarity 100.0%; Pred. No. 0.57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KNEYTVHL 9
Db 1 KNEYTVHL 8

RESULT 3
Q9G0B3 PRELIMINARY; PRT; 210 AA.
AC Q9G0B3;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Distal tail fiber locus, partial sequence (Fragment).
GN T.
OS Bacteriophage AR1.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC T4-like viruses.
OX NCBI_TaxID=66711;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20485545; PubMed=11029414;
RA Yu S.L., Ko K.B., Chen C.S., Chang Y.C., Syu W.J.;
RT "Characterization of the Distal Tail Fiber Locus and Determination of
the Receptor for Phage AR1, Which Specifically Infects Escherichia
coli O157:H7";
RL J. Bacteriol. 182:5962-5968 (2000).
DR EMBL; AF208841; AAG29756.1; -.
FT NON_TER 210 210

SQ SEQUENCE 210 AA; 24348 MW; C6FF585F9FE0DF68 CRC64;
Query Match 79.6%; Score 39; DB 9; Length 210;
Best Local Similarity 87.5%; Pred. No. 7;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 KNEYTVHL 9
Db 139 KNEYTVHL 146

RESULT 4
Q8DPG2 PRELIMINARY; PRT; 334 AA.
AC Q8DPG2;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Conserved hypothetical protein.
GN SPRI179.
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=171101;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21429245; PubMed=11544234;
RA Hoskins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
DeHoff B.S., Batrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McAhren S.M., McHenry M., Mcleaster K., Mundy C.W., Nicae T.I.,
RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rocky P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,
RA Glass J.I.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6";
RL J. Bacteriol. 183:5709-5717 (2001).
DR EMBL; AB008490; AAK9982.1; -.
DR InterPro; IPR006314; Dyp_peroxidase.
DR Pfam; PF04261; Dyp_perox; 1.
DR TIGRFAMs; tigr01413; Dyp_perox_fam; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 334 AA; 38137 MW; E207F1BC267334E5 CRC64;

Query Match 75.5%; Score 37; DB 16; Length 334;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KNEYTVHL 9
Db 302 KNEYTVHL 310

RESULT 5
Q8L836 PRELIMINARY; PRT; 603 AA.
AC Q8L836;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein (At1g14690).
GN AT1G14690.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Tripp M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,

RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
 RA Ecker J.R., Theologis A., Davis R.W.,
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
 RA Dale J.M., Hayashizaki Y., Huan V.W., Ishida J., Jones T., Kamiya A.,
 RA Karlin-Neumann G., Kawai J., Lam B., Lin J., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
 RA Seki M., Southwick A., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G.,
 RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.,
 RT "Arabidopsis ORF clones",
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 DR ENBL; AY120768; AAM53326.1; -
 DR ENBL; BT008373; AAP37732.1; -
 DR InterPro: IPR007145; MAP65_ASE1.
 DR Pfam; PF03999; MAP65_ASE1; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 603 AA; 69061 MW; 7C2B37A23CB889D CRC64;

Query Match 73.5%; Score 36; DB 10; Length 603;
 Best Local Similarity 66.7%; Pred. NO. 82;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KMNEYTVHL 9
 Db 169 KLNQYQTHL 177

RESULT 5

ID Q9XU04 PRELIMINARY; PRT; 627 AA.
 AC Q9XU04;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE T28A8.6 protein.
 GN T28A8.6.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lloyd C.R.,
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RL Science 282:2012-2018 (1998).
 DR EMBL; Z92813; CAB07289.1; -
 DR PIR; T25395; T25395.
 DR WormPep; T28A8.6; CE18977.
 DR InterPro: IPR006570; SPK.
 DR Pfam; PF04435; SPK; 2.
 DR SMART; SM00583; SPK; 2.
 SQ SEQUENCE 627 AA; 71794 MW; 4A56E4D36585C1CC CRC64;

Query Match 73.5%; Score 36; DB 5; Length 627;
 Best Local Similarity 66.7%; Pred. NO. 86;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KMNEYTVHL 9
 Db 62 KMNEYSIEL 70

RESULT 7

Q9LQW1
 ID Q9LQW1 PRELIMINARY; PRT; 635 AA.
 AC Q9LQW1;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE F10B6.9.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chao Q., Shinn P., Dunn P., Buehler E., Kahn S., Kim C., Walker M.,
 RA Williams S., Altafi H., Araujo R., Conn L., Conway A.B., Gonzalez A.,
 RA Hansen N.F., Huizar L., Kremenetskaia I., Lenz C., Li J., Liu S.,
 RA Luros S., Rowley D., Schwartz J., Toriumi M., Vysotskaia V., Yu G.,
 RA Davis R.W., Federspiel N.A., Theologis A., Ecker J.R.,
 RT "Genomic sequence for Arabidopsis thaliana BAC F10B6 from chromosome
 RT 1",
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.,
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.,
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
 RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E.,
 RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
 RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsy N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
 RA Theologis A., Ecker J.,
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC006917; AAF79248.1; -
 DR InterPro: IPR007145; MAP65_ASE1.
 DR Pfam; PF03999; MAP65_ASE1; 1.
 SQ SEQUENCE 635 AA; 72664 MW; 2871C7DA1E4DA65E CRC64;

Query Match 73.5%; Score 36; DB 10; Length 635;
 Best Local Similarity 66.7%; Pred. NO. 87;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KMNEYTVHL 9
 Db 169 KLNQYQTHL 177

RESULT 8

ID Q8E6M8 PRELIMINARY; PRT; 753 AA.
 AC Q8E6M8;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN GBS0535.
 OS Streptococcus agalactiae (serotype III).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=216495;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NEM316 / Serotype III;
 RX MEDLINE=2242508; PubMed=12354221;
 RA Glaser P., Rusniok C., Suchrieser C., Chevalier F., Frangeul L.,
 RA Meadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trier-Cuot P.,

Kunst F.;
"genome sequence of Streptococcus agalactiae, a pathogen causing
invasive neonatal disease."; Mol. Microbiol. 45:1499-1513(2002).
EMBL; AL766845; CAD46179.1; -.
Sagalist; gbs0535; -.
GO; GO:0005524; F:ATP binding; IEA.
GO; GO:0003677; F:DNA binding; IEA.
GO; GO:0004518; F:nuclease activity; IEA.
GO; GO:0000166; F:nucleotide binding; IEA.
GO; GO:0006289; F:nucleotide-excision repair; IEA.
InterPro; IPR003593; AAA_ATPase.
InterPro; IPR001270; Chaprinin_clpA/B.
InterPro; IPR001943; UvrB/C.
Pfam; PF00004; AAA; 2.
Pfam; PF02151; UVR; 1.
PRINTS; PR00300; CLPTEASEA.
SMART; SMO0362; AAA; 2.
PROSITE; PS00870; CLPAB_1; 1.
PROSITE; PS00871; CLPAB_2; 1.
PROSITE; PSS0151; UVR; 1.
Hypothetical protein; Complete proteome.
SEQUENCE 753 AA; 83803 MW; 5DFOB1230E47555D CRC64;

Query Match 73.5%; Score 36; DB 16; Length 753;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0;

Qv 1 KMNEYTVHL 9
|:|:|:|
Db 7 KLNESIHL 15

RESULT 9

QSE174 PRELIMINARY; PRT; 753 AA.

AC Q8E174
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ATP-dependent Clp protease, ATP-binding subunit.
GN SAGC488
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
CX NCBI_TaxID=216466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=2222988; PubMed=12200547;
RA Wessells H., Maignan V., Cieslewicz M.J., Eisen J.A., Peterson
RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.W., Daugherty S.C.,
RA DeBoy R.T., Purkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
RA Carthy H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mc
RA Iacobino E.T., Brettoni C., Galli G., Mariani M., Vegni F., Mai
RA Rinaldo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA Praser C.W.;
RT "Complete genome sequence and comparative genomic analysis of
emerging human pathogen, serotype V Streptococcus agalactiae";
Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
RL EMBL; AE014214; AAM99390.1; -.
DR TIGR; SAG0488; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004519; F:nuclease activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006283; F:peptidase activity; IEA.
DR GO; GO:0006289; F:nucleotide-excision repair; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase.

```

RESULT 11
Q19815 PRELIMINARY; PRT; 1829 AA.
AC Q19815;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE F26F12.7 protein (LET-418).
GN F26F12.7 OR LET-418.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
SEQUENCE FROM N.A.
RP STRAIN=BRISTOL N2;
RC MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
RA Smalton N., Smith A., Sonnenhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38 (1994).
[2]
SEQUENCE FROM N.A.
RP STRAIN=BRISTOL N2;
RC STRAIN=BRISTOL N2;
RA Wilson R., Bentley D., Gattung S.;
RT "The sequence of C. elegans cosmid F26F12.";
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
RP STRAIN=BRISTOL N2;
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE FROM N.A.
RP STRAIN=BRISTOL N2;
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
[5]
SEQUENCE FROM N.A.
RP von Zelewsky T., Palladino F., Brunschwig K., Hajnal A., Mueller F.;
RA "The C. elegans Mi-2 chromatin-remodeling proteins function in vulval
RT cell fate determination.";
RL Development 0:0-0(2000).
DR EMBL; U55373; AAC25894.1; -.
DR EMBL; AF308445; AAG29838.1; -.
DR PIR; T34239; T34239.
DR WormPep; F26F12.7; Csl7716.
DR GO; GO:000785; C:chromatin; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP dependent helicase activity; IEA.
DR GO; GO:0003682; F:chromatin binding; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0006333; P:chromatin assembly/disassembly; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000953; Chromo.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002464; DEAD box.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR000330; SNF2_N.
DR InterPro; IPR001985; Znf_FHD.
DR InterPro; IPR001841; Znf_Fing.
PFam; PF00385; chromo; 1.

```

```

DR Pfam; PF00271; helicase_C; 1.
DR Pfam; PF00628; PHD; 2.
DR Pfam; PF00176; SNF2_N; 1.
DR SMART; SM00298; CHROMO; 2.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICc; 1.
DR SMART; SM00249; PHD; 2.
DR SMART; SM00184; RING; 2.
DR PROSITE; PS00013; CHROMO 2; 2.
DR PROSITE; PS00690; DEAD_ATP_HELICASE; 1.
DR PROSITE; PS00016; ZF_PHD 2; 2.
KW ATP-binding; Helicase; Hydrolase.
SQ SEQUENCE 1829 AA; 209150 MW; 1A887E990C63B661 CRC64;

Query Match 73.5%; Score 36; DB 5; Length 1829;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MNEYTVH 8
Db 765 LNEYTIH 771

RESULT 12
Q810Y2 PRELIMINARY; PRT; 47 AA.
AC Q810Y2;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE FYI protein (Fragment).
GN MID1.
OS Mus spicilegus (Steppe mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZRU;
RX MEDLINE=22503854; PubMed=12615004;
RA Montoya-Burgos J.I., Boursot P., Galtier N.;
RT "Recombination explains isochores in mammalian genomes.";
RL Trends Genet. 19:128-130(2003).
DR EMBL; AY181235; AAG72138.1; -.
DR InterPro; IPR008957; FN_III-like.
FT NON_TER 1
FT NON_TER 47
FT NON_TER 47
SQ SEQUENCE 47 AA; 5261 MW; 9DC0A9F99DBCC1A6 CRC64;

Query Match 71.4%; Score 35; DB 11; Length 47;
Best Local Similarity 75.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KQNEYTVH 8
Db 9 KQNEYTVH 16

RESULT 13
Q80ZB4 PRELIMINARY; PRT; 47 AA.
AC Q80ZB4;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE FYI (Fragment).
OS Mus platyrrhinus (Flat-haired mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10101;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PTX;

```

RX MEDLINE=22503854; PubMed=12615004;
 RA Montoya-Burgos J.I., Boursot P., Galtier N.;
 RT "Recombination explains isochores in mammalian genomes."
 RL Trends Genet. 19:128-130(2003).
 DR EMBL; AY181223; AAC62980.1; -;
 DR InterPro; IPR008957; FN_III-like.
 FT NON_TER 1
 FT NON_TER 47
 SQ SEQUENCE 47 AA; 5261 MW; 9DC0A9F99DBCC1A6 CRC64;

Query Match 71.4%; Score 35; DB 11; Length 47;
 Best Local Similarity 75.0%; Pred. No. 10;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KXNEYTVH 8
 DB 9 KQHHTVH 16

RESULT 14

Q80283 PRELIMINARY; PRT; 47 AA.
 AC Q80283;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Fxy protein (Fragment).
 GN FXY.
 OS Arvicola terrestris (European water vole).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Arvicolinae;
 OC Arvicolinae;
 OC NCBI_TaxID=10050;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22503854; PubMed=12615004;
 RA Montoya-Burgos J.I., Boursot P., Galtier N.;
 RT "Recombination explains isochores in mammalian genomes."
 RL Trends Genet. 19:128-130(2003).
 DR EMBL; AY181220; AAC62980.1; -;
 DR InterPro; IPR008957; FN_III-like.
 FT NON_TER 1
 FT NON_TER 47
 SQ SEQUENCE 47 AA; 5261 MW; 9DC0A9F99DBCC1A6 CRC64;

Query Match 71.4%; Score 35; DB 11; Length 47;
 Best Local Similarity 75.0%; Pred. No. 10;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KXNEYTVH 8
 DB 9 KQHHTVH 16

RESULT 15

Q80280 PRELIMINARY; PRT; 47 AA.
 AC Q80280;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Fxy protein (Fragment).
 GN FXY.
 OS Mastomys huberti.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 OC Mastomys.
 OC NCBI_TaxID=121569;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22503854; PubMed=12615004;
 RA Montoya-Burgos J.I., Boursot P., Galtier N.;
 RT "Recombination explains isochores in mammalian genomes."

RL Trends Genet. 19:128-130(2003).
 DR EMBL; AY181223; AAC62983.1; -;
 DR InterPro; IPR008957; FN_III-like.
 FT NON_TER 1
 FT NON_TER 47
 SQ SEQUENCE 47 AA; 5261 MW; 9DC0A9F99DBCC1A6 CRC64;

Query Match 71.4%; Score 35; DB 11; Length 47;
 Best Local Similarity 75.0%; Pred. No. 10;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KXNEYTVH 8
 DB 9 KQHHTVH 16

Search completed: March 1, 2004, 17:34:39
 Job time : 32.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:21:46 ; Search time 11.8899 Seconds
(without alignments)
39.081 Million cell updates/sec

Title: US-09-905-083-32
Perfect score: 40
Sequence: 1 RLSSMVKKV 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	9	3	US-09-502-600-32
2	40	100.0	9	4	US-09-918-243-32
3	40	100.0	144	4	US-09-618-259-4
4	40	100.0	154	3	US-09-261-416-7
5	40	100.0	161	3	US-09-261-416-8
6	40	100.0	224	3	US-08-944-483-33
7	40	100.0	225	2	US-08-557-146-12
8	40	100.0	225	2	US-09-027-337-4
9	40	100.0	225	2	US-09-154-344-12
10	40	100.0	225	4	US-09-644-600-4
11	40	100.0	225	4	US-09-654-600A-4
12	40	100.0	253	2	US-08-557-146-2
13	40	100.0	253	2	US-08-824-874-3
14	40	100.0	253	2	US-09-154-344-2
15	40	100.0	253	3	US-08-930-188-2
16	40	100.0	253	3	US-09-210-084-3
17	40	100.0	253	4	US-09-764-762-3
18	40	100.0	253	5	PCT-US96-04294-2
19	36	90.0	9	3	US-09-502-600-110
20	36	90.0	9	4	US-09-918-243-110
21	31	77.5	9	3	US-09-502-600-79
22	31	77.5	9	3	US-09-502-600-87
23	31	77.5	9	4	US-09-918-243-79
24	31	77.5	9	4	US-09-918-243-87
25	30	75.0	156	4	US-09-328-352-7506
26	29	72.5	571	4	US-09-690-942-15
27	29	72.5	582	1	US-08-261-086-2

28 72.5 582 1 US-08-261-086-4 Sequence 4, Appli
29 72.5 582 1 US-08-261-086-6 Sequence 6, Appli
30 154 4 US-09-328-352-7990 Sequence 7990, Ap
31 28 70.0 250 4 US-09-107-532A-4737 Sequence 4737, Ap
32 28 70.0 289 4 US-09-107-532A-6154 Sequence 6154, Ap
33 28 70.0 341 4 US-09-134-001C-3650 Sequence 3650, Ap
34 28 70.0 405 4 US-09-134-000C-5465 Sequence 5465, Ap
35 28 70.0 459 4 US-09-491-785-2 Sequence 2, Appli
36 28 70.0 464 4 US-09-134-001C-4562 Sequence 4562, Ap
37 28 70.0 565 6 5171685-6 Patent No. 5171685
38 28 70.0 565 6 5171685-6 Patent No. 5171685
39 28 70.0 765 4 US-09-134-001C-3681 Sequence 3681, Ap
40 28 70.0 861 4 US-09-107-532A-5985 Sequence 5985, Ap
41 28 70.0 883 4 US-08-982-430-1 Sequence 1, Appli
42 28 70.0 896 2 US-08-640-389A-10 Sequence 10, Appl
43 28 70.0 896 4 US-08-618-957A-10 Sequence 10, Appl
44 28 70.0 896 4 US-09-043-816E-13 Sequence 13, Appl
45 28 70.0 896 4 US-09-357-914-33 Sequence 33, Appl

ALIGNMENTS

RESULT 1
US-09-502-600-32
; Sequence 32, Application US/09502600A
; Patent No. 6294344
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP-C
; CURRENT FILING DATE: 2000-02-11
; CURRENT APPLICATION NUMBER: US/09/502,600A
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 32
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 123-131 of the SCCE protein
US-09-502-600-32

Query Match 100.0%; Score 40; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05; 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0;

Qy 1 RLSSMVKKV 9
|||
Db 1 RLSSMVKKV 9

RESULT 2
US-09-918-243-32
; Sequence 32, Application US/09918243
; Patent No. 6627403
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 32
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens

```
FEATURE:
NAME/KEY: CHAIN
OTHER INFORMATION: Residues 123-131 of the SCCE protein
US-09-918-243-32

Query Match
Best Local Similarity 100.0%; Score 40; DB 4; Length 9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9
Db 1 RLSSMVKKV 9

RESULT 3
US-09-618-259-4
Sequence 4, Application US/09618259
Patent No. 6642013
GENERAL INFORMATION:
APPLICANT: O'Brien, Timothy J.
APPLICANT: Underwood, Lowell J.
TITLE OF INVENTION: No. 6642013el Extracellular Serine Protease
FILE REFERENCE: D6020CIP2
CURRENT APPLICATION NUMBER: US/09/618,259
CURRENT FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: US 09/127,444
PRIOR FILING DATE: 1998-08-21
NUMBER OF SEQ ID NOS: 72
SEQ ID NO 4
LENGTH: 144
TYPE: PRT
ORGANISM: unknown
FEATURE:
NAME/KEY: DOMAIN
OTHER INFORMATION: Amino acid sequence of stratum corneum chymotryptic
OTHER INFORMATION: enzyme (sccc) catalytic domain
US-09-618-259-4

Query Match
Best Local Similarity 100.0%; Score 40; DB 4; Length 144;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9
Db 60 RLSSMVKKV 68

RESULT 4
US-09-261-416-7
Sequence 7, Application US/09261416A
Patent No. 6291663
GENERAL INFORMATION:
APPLICANT: O'Brien, Timothy J.
APPLICANT: Underwood, Lowell J.
TITLE OF INVENTION: TADG-12: A No. 6291663el Transmembrane Serine Protease
FILE REFERENCE: D6192
CURRENT APPLICATION NUMBER: US/09/261,416A
CURRENT FILING DATE: 1999-03-03
NUMBER OF SEQ ID NOS: 14
SEQ ID NO 7
LENGTH: 154
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Serine protease catalytic domain of stratum corneum
OTHER INFORMATION: chymotryptic enzyme (Sccc) homologous to similar
OTHER INFORMATION: domain in TADG-12
US-09-261-416-7

Query Match
Best Local Similarity 100.0%; Score 40; DB 3; Length 154;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9
Db 69 RLSSMVKKV 77

RESULT 5
US-09-261-416-8
Sequence 8, Application US/09261416A
Patent No. 6291663
GENERAL INFORMATION:
APPLICANT: O'Brien, Timothy J.
APPLICANT: Underwood, Lowell J.
TITLE OF INVENTION: TADG-12: A No. 6291663el Transmembrane Serine Protease
FILE REFERENCE: D6192
CURRENT APPLICATION NUMBER: US/09/261,416A
CURRENT FILING DATE: 1999-03-03
NUMBER OF SEQ ID NOS: 14
SEQ ID NO 8
LENGTH: 161
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Serine protease catalytic domain of hepsin (Heps)
OTHER INFORMATION: homologous to similar domain in TADG-12
US-09-261-416-8

Query Match
Best Local Similarity 100.0%; Score 40; DB 3; Length 161;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9
Db 75 RLSSMVKKV 83

RESULT 6
US-08-344-483-33
Sequence 33, Application US/08344483
Patent No. 6232456
GENERAL INFORMATION:
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: KLASS, MICHAEL R.
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STEWART, KENT D.
APPLICANT: STROUPE, STEVEN D.
TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/944,483
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
```


ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6183. US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 224 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6232456e
US-08-944-483-33

Query Match 100.0%; Score 40; DB 3; Length 224;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9
Db 94 RLSSMVKKV 102

RESULT 7
US-08-557-146-12
Sequence 12, Application US/08557145
Patent No. 5834290
GENERAL INFORMATION:
APPLICANT: Egelrud, Torbjorn
APPLICANT: Hansson, Lennart
TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
TITLE OF INVENTION: Enzyme (SCCE)
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: White & Case, Patent Department
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/557,146
FILING DATE: 14-DEC-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sterner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-181
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 225 amino acids
TYPE: amino acids
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: polypeptide
US-08-557-146-12

Query Match 100.0%; Score 40; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9
Db 94 RLSSMVKKV 102

RESULT 8
US-09-027-337-4
Sequence 4, Application US/09027337B
Patent No. 5972616
GENERAL INFORMATION:
APPLICANT: O'Brien, Timothy J.
APPLICANT: Tanimoto, Hirotochi
TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease Overexpressed in
TITLE OF INVENTION: Breast and Ovarian Carcinomas
FILE REFERENCE: D6064
CURRENT APPLICATION NUMBER: US/09/027,337B
CURRENT FILING DATE: 1998-02-20
NUMBER OF SEQ ID NOS: 13
SEQ ID NO 4
LENGTH: 225
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Serine protease catalytic domain of Ssce homologous to
OTHER INFORMATION: similar domain in TADG-15
US-09-027-337-4

Query Match 100.0%; Score 40; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9
Db 95 RLSSMVKKV 103

RESULT 9
US-09-154-344-12
Sequence 12, Application US/09154344
Patent No. 5981256
GENERAL INFORMATION:
APPLICANT: Egelrud, Torbjorn
APPLICANT: Hansson, Lennart
TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
TITLE OF INVENTION: Enzyme (SCCE)
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: White & Case, Patent Department
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/154,344
FILING DATE: 16-SEP-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/557,146
FILING DATE: 14-DEC-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sterner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-181
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 12:

```
/
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 225 amino acids
/ TYPE: amino acids
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: polypeptide
US-09-154-344-12

Query Match 100.0%; Score 40; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9
Db 95 RLSSMVKKV 103

RESULT 10
US-09-644-600-4
; Sequence 4, Application US/09644600
; Patent No. 6451500
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotoishi
; TITLE OF INVENTION: TAGD-15: An Extracellular Serine Protease
; FILE OF INVENTION: Overexpressed in Carcinomas
; FILE REFERENCE: D6064CIP/D
; CURRENT APPLICATION NUMBER: US/09/644,600
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/421,213
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: 09/027,337
; PRIOR FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 4
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: SCCE
US-09-644-600-4

Query Match 100.0%; Score 40; DB 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9
Db 95 RLSSMVKKV 103

RESULT 11
US-09-654-600A-4
; Sequence 4, Application US/09654600A
; Patent No. 664974
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotoishi
; TITLE OF INVENTION: TAGD-15: An Extracellular Serine Protease
; FILE OF INVENTION: Overexpressed in Carcinomas
; FILE REFERENCE: D6064CIP/D
; CURRENT APPLICATION NUMBER: US/09/654,600A
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/421,213
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 4
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: SCCE
US-09-654-600A-4

Query Match 100.0%; Score 40; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9
Db 123 RLSSMVKKV 131

RESULT 12
US-08-557-146-2
; Sequence 2, Application US/08557146
; Patent No. 5834290
; GENERAL INFORMATION:
; APPLICANT: Egeirud, Torbjorn
; APPLICANT: Hansson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,146
; FILING DATE: 14-DEC-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-557-146-2

Query Match 100.0%; Score 40; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9
Db 123 RLSSMVKKV 131

RESULT 13
US-08-824-874-3
; Sequence 3, Application US/08824874
; Patent No. 5952300
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NOVEL KALLIKREIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; OTHER INFORMATION: SCCE
US-08-824-874-3
```

```

/ STREET: 3174 Porter Drive
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FASTSEQ for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/824,874
/ FILING DATE: Filed Herewith
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0252 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-853-0555
/ TELEFAX: 415-845-4166
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 253 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GenBank
/ CLONE: 532504
/ US-08-824-874-3

Query Match 100.0%; Score 40; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9
Db 123 RLSSMVKKV 131

RESULT 14
US-09-154-344-2
/ Sequence 2, Application US/09154344
/ Patent No. 5981256
/ GENERAL INFORMATION:
/ APPLICANT: Egelrud, Torbjorn
/ APPLICANT: Hagason, Lennart
/ TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
/ NUMBER OF SEQUENCES: 17
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: White & Case, Patent Department
/ STREET: 1155 Avenue of the Americas
/ CITY: New York
/ STATE: New York
/ COUNTRY: U.S.A.
/ ZIP: 10036-2787
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/154,344
/ FILING DATE: 16-SEP-1998
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/557,146
/ FILING DATE: 14-DEC-1995

```

```

/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Sterner, Richard J.
/ REGISTRATION NUMBER: 35,372
/ REFERENCE/DOCKET NUMBER: 1103326-181
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 819-8783
/ TELEFAX: (212) 354-8113
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 253 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-09-154-344-2

Query Match 100.0%; Score 40; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9
Db 123 RLSSMVKKV 131

RESULT 15
US-08-930-188-2
/ Sequence 2, Application US/08930188
/ Patent No. 6093397
/ GENERAL INFORMATION:
/ APPLICANT: Dixon, Eric P.
/ APPLICANT: Johnstone, Edward M.
/ APPLICANT: Little, Sheila P.
/ TITLE OF INVENTION: AMYLOID PRECURSOR PROTEIN PROTEASE AND
/ TITLE OF INVENTION: RELATED NUCLEIC ACIDS
/ NUMBER OF SEQUENCES: 3
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Eli Lilly and Company
/ STREET: Lilly Corporate Center
/ CITY: Indianapolis
/ STATE: Indiana
/ COUNTRY: United States of America
/ ZIP: 46285
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/930,188
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/416,257
/ FILING DATE: 04-APR-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Blalock, Donna K.
/ REGISTRATION NUMBER: 38,082
/ REFERENCE/DOCKET NUMBER: X9239
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 317-277-1090
/ TELEFAX: 317-276-3861
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 253 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-930-188-2

Query Match 100.0%; Score 40; DB 3; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.2;

```

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMVKKV 9

|||||

Db 123 RLSSMVKKV 131

Search completed: March 1, 2004, 17:38:23
Job time : 12.8889 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:16:55 ; Search time 45.5556 Seconds
(without alignments)
55.820 Million cell updates/sec

Title: US-09-905-083-32
Perfect score: 40
Sequence: 1 RLSSMVKKV 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A Geneseq_29Jan04.*
- 1: geneseqp1980s.*
 - 2: geneseqp1980s.*
 - 3: geneseqp2000s.*
 - 4: geneseqp2000s.*
 - 5: geneseqp2002s.*
 - 6: geneseqp2003as.*
 - 7: geneseqp2003bs.*
 - 8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	40	100.0	9	4	Aae08237 Human str
2	40	100.0	224	6	ADA05744 Human NOV
3	40	100.0	225	4	AAB98502 Human Str
4	40	100.0	247	6	ADA05742 Human NOV
5	40	100.0	250	6	ADA05732 Human NOV
6	40	100.0	252	6	ADA05734 Human NOV
7	40	100.0	253	2	AAR67888 Human str
8	40	100.0	253	2	AAR05383 Human am
9	40	100.0	253	5	ABB84421 Human SCC
10	40	100.0	253	5	ABB84406 Human SCC
11	40	100.0	253	5	AAB82740 Amino aci
12	40	100.0	253	6	ABU07440 Protein d
13	40	100.0	253	6	ABU07471 Protein d
14	40	100.0	253	6	ABR58471 Human str
15	40	100.0	253	6	ADB80484 Ovarian c
16	40	100.0	257	3	AAB21326 Human HSC
17	36	90.0	9	4	AAE08314 Human str
18	33	82.5	394	5	ABP52834 Anopheles
19	33	82.5	394	7	ABR84477 Mosquito
20	33	82.5	417	7	ADD15250 African m
21	32	80.0	589	6	ABU48685 Protein e
22	31	77.5	9	4	AAE08292 Human str
23	31	77.5	9	4	AAE08284 Human str
24	31	77.5	136	4	ABG22378 Novel hum
25	31	77.5	229	3	AAG19496 Arabidops

26 31 77.5 229 3 AAG52872 Arabidops
27 31 77.5 243 5 ABB84419 Bovine SC
28 31 77.5 249 5 ABB84420 Porcine S
29 31 77.5 390 3 AAG19495 Arabidops
30 31 77.5 390 3 AAG52871 Arabidops
31 31 77.5 394 3 AAG52870 Arabidops
32 31 77.5 394 3 AAG19494 Arabidops
33 31 77.5 2159 6 ABP70924 Maize DEK
34 30 75.0 156 6 ADA36219 Acinetoba
35 30 75.0 220 3 AAB24712 Plant SDF
36 30 75.0 220 3 AAB25046 Plant SDF
37 30 75.0 220 3 AAG38343 Arabidops
38 30 75.0 221 3 AAG38345 Arabidops
39 30 75.0 236 3 AAB5045 Arabidops
40 30 75.0 236 3 AAB24711 Plant SDF
41 30 75.0 236 3 AAG38342 Arabidops
42 30 75.0 237 3 AAG38344 Arabidops
43 30 75.0 318 7 ADD48786 Rat Prote
44 30 75.0 318 7 ADE62226 Rat Prote
45 30 75.0 447 6 ABU44337 Protein e

ALIGNMENTS

RESULT 1
AAE08237
ID AAE08237 standard; peptide; 9 AA.

XX AAE08237;
XX AC
XX 01-NOV-2001 (first entry)
XX Human stratum corneum chymotrypsin enzyme peptide #2 (residues 123-131).
XX Stratum corneum chymotrypsin enzyme; SCCE; cytostatic; vaccine; tumour;
XX cancer; ovarian; breast; lung; colon; prostate; carcinoma; human;
XX antitense therapy; malignant hyperplasia.
XX Homo sapiens.
XX WO200159158-A1.
XX 16-AUG-2001.
XX 07-FEB-2001; 2001WO-US003977.
XX 11-FEB-2000; 2000US-00502600.
XX (UYAR-) UNIV ARKANSAS.
XX O'brien TJ;
XX WPI; 2001-514676/56.
XX Diagnosing cancer comprises detecting stratum corneum chymotrypsin enzyme.

Claim 25; Page 102; 127pp; English.
The invention relates to diagnosing cancer especially ovarian cancer, by screening for stratum corneum chymotrypsin enzyme (SCCE). Proteases are considered to be an integral part of tumour growth and metastasis, and therefore, markers indicative of their presence or absence are useful for the diagnosis of cancer. The method is useful for diagnosing cancer, carcinoma and malignant hyperplasia. The SCCE oligonucleotide may be used to treat a cancer selected from ovarian, breast, lung, colon, prostate and other cancers in which SCCE is overexpressed. The present sequence is human SCCE peptide

XX Sequence 9 AA;

Query Match 100.0%; Score 40; DB 4; Length 9;

Best Local Similarity 100.0%, Pred. NO. 1.4e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RLSSMVKKV 9
|||||

Db 1 RLSSMVKKV 9

RESULT 2
ADA05744
IDA05744 standard; protein; 224 AA.
XX AC
ADA05744;
XX XX
DT 06-NOV-2003 (first entry)
XX XX
DE Human NOV18g protein SEQ ID NO:104.
XX XX
KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
KW immunomodulator; cytostatic; nootropic; neuroprotective;
KW antiparkinsonian; antilipase; gene therapy; human disease;
KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW immune disorder; haematopoietic disorder; dyslipidaemia.
XX XX
OS Homo sapiens.
XX XX
FN WO2003029424-A2.
XX XX
PD 10-APR-2003.
XX XX
PF 02-OCT-2002; 2002WO-US031373.
XX XX
PR 02-OCT-2001; 2001US-0326483P.
PR 03-OCT-2001; 2001US-0327435P.
PR 03-OCT-2001; 2001US-0327449P.
PR 09-OCT-2001; 2001US-0327917P.
PR 09-OCT-2001; 2001US-0328029P.
PR 09-OCT-2001; 2001US-0328044P.
PR 09-OCT-2001; 2001US-0328056P.
PR 12-OCT-2001; 2001US-0328849P.
PR 15-OCT-2001; 2001US-0329414P.
PR 17-OCT-2001; 2001US-0330142P.
PR 18-OCT-2001; 2001US-0330309P.
PR 22-OCT-2001; 2001US-0341058P.
PR 24-OCT-2001; 2001US-0339266P.
PR 24-OCT-2001; 2001US-0343629P.
PR 29-OCT-2001; 2001US-0349575P.
PR 01-NOV-2001; 2001US-0346357P.
PR 17-APR-2002; 2002US-0373260P.
PR 19-APR-2002; 2002US-0373815P.
PR 19-APR-2002; 2002US-0373817P.
PR 19-APR-2002; 2002US-0373826P.
PR 19-APR-2002; 2002US-0373884P.
PR 22-APR-2002; 2002US-0374977P.
PR 16-MAY-2002; 2002US-0381037P.
PR 16-MAY-2002; 2002US-0381038P.
PR 16-MAY-2002; 2002US-0381042P.
PR 17-MAY-2002; 2002US-0381642P.
PR 28-MAY-2002; 2002US-0383656P.
PR 29-MAY-2002; 2002US-0383831P.
PR 25-JUN-2002; 2002US-0391335P.
PR 01-OCT-2002; 2002US-00262511.
XX XX
(CURA-) CURAGEN CORP.
XX XX
PI Smithson G, Mallett I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
PI Paturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
PI Ott T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;
PI Ji W, Maller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;
PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;
PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;
XX XX

WPI; 2003-391626/36.
N-PSDB; ADA05743.

New NOVX polypeptides and nucleic acids, useful for diagnosing, preventing or treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or dyslipidemia, and in chromosome mapping, tissue typing or pharmacogenomics.

Claim 1; Page 172; 586pp; English.

The present invention describes NOVX proteins, where X can be l to 55 (e.g. NOVl). Also described: (1) a composition comprising a polypeptide described above and a carrier; (2) a kit comprising, in one or more containers, the composition described above; (3) an isolated nucleic acid molecule which encodes a NOVX protein of the invention; (4) a vector comprising the nucleic acid molecule described above; (5) a cell comprising the above vector; (6) an antibody that immunospecifically binds to the polypeptide described above; (7) methods for determining the presence or amount of the above polypeptide or nucleic acid molecule in a sample; (8) methods for determining the presence of or predisposition to a disease associated with altered levels of expression of the above polypeptide or nucleic acid molecule in a first mammalian subject; (9) a method of identifying an agent that binds to the polypeptide described above; (10) a method for identifying a potential therapeutic agent for use in treating a pathology that is related to an aberrant expression or aberrant physiological interactions of the polypeptide; (11) a method of screening for a modulator of activity or of latency or predisposition to a pathology associated with the polypeptide; (12) a method for modulating the activity of the polypeptide described above; (13) methods of treating or preventing a pathology associated with the above polypeptide in a mammal; and (14) a method for producing the above polypeptide. NOVX sequences have antidiabetic, anorectic, antibacterial, virucide, immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian and anti-lipidemic activities, and can be used in gene therapy. The polypeptide is useful in manufacturing a medicament for treating a syndrome associated with a human disease. The polypeptide or the nucleic acid molecule may be used to diagnose, treat or prevent metabolic disorders such as diabetes or obesity, infections, cachexia, cancer, neurodegenerative disorders such as Alzheimer's disease or Parkinson's disease, immune disorders, haematopoietic disorders and various dyslipidaemias. The nucleic acids can also be used as hybridisation probes, in chromosome mapping, tissue typing, preventive medicine and pharmacogenomics. The present sequence represents a human NOVX from the present invention.

Sequence 224 AA;

Query Match 100.0%; Score 40; DB 6; Length 224;
Best Local Similarity 100.0%; P-red. No. 2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RLSSMWKKV 9
| | | | |
Db 109 RLSSMWKKV 117

RESULT 3
AAB98502
ID AAB98502 standard; protein; 225 AA.
XX AAB98502;
AC AAB98502;
XX
XX 03-AUG-2001 (first entry)
XX Human Stratum Corneum Chymotryptic Enzyme, SCCE, catalytic domain.
XX Human; TADG-15; cytosstatic; vaccine; ovarian tumour; cancer;
KW tumour antigen-derived gene 15; serine protease;
KW Stratum Corneum Chymotryptic Enzyme; SCCE.
XX Homo sapiens.
OS
XX WO200129056-A1.

XX PD 26-APR-2001.
XX PF 20-OCT-2000; 2000WO-US029095.
XX PR 20-OCT-1999; 99US-00421213.
XX PR (UYAR-) UNIV ARKANSAS.
XX PA O'brien TJ, Tanimoto H;
PI WPI; 2001-381031/40.
DR XX
XX Novel extracellular serine protease, termed tumor antigen-derived gene 15
PT protein overexpressed in carcinomas and DNA encoding it, for diagnosis,
PT treatment, prevention of cancer, particularly breast, ovarian cancer.
XX Example 10; Fig 1; 130pp; English.
PS
XX The present invention relates to human tumour antigen-derived gene 15
CC (TADG-15) protein and coding sequence (see AAH23601 and AAB98500). TADG-
CC 15 is an extracellular serine protease. It was found that TADG-15 is over
CC -expressed in ovarian tumours. TADG-15 protein or its fragments of 9-20
CC residues that lack TADG-15 protease activity are useful for vaccinating
CC an individual against TADG-15, having suspected of having or at risk of
CC getting cancer. Furthermore, the TADG-15 gene can be used as a diagnostic
CC or therapeutic target in cancer. The present sequence was used in a
CC sequence homology alignment with the catalytic domain of TADG-15
XX
XX Sequence 225 AA;
Query Match 100.0%; Score 40; DS 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 RLSSMVKKV 9
Db 95 RLSSMVKKV 103
|||||
|
RESULT 4
ADA05742
ID ADA05742 standard; protein; 247 AA.
XX AC ADA05742;
XX DT 06-NOV-2003 (first entry)
XX DE Human NOV18f protein SEQ ID NO:102.
XX KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
KW immunomodulator; cytostatic; neurotropic; neuroprotective;
KW antiparkinsonian; antilipemic; gene therapy; human disease;
KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
KW neurodegenerative disorder; Alzheimer's disease; parkinson's disease;
KW immune disorder; haematopoietic disorder; dyslipidaemia.
XX OS Homo sapiens.
XX PN WO2003029424-A2.
XX PD 10-APR-2003.
XX PF 02-OCT-2002; 2002WO-US031373.
XX PR 02-OCT-2001; 2001US-0326483P.
XX PR 05-OCT-2001; 2001US-0327435P.
XX PR 09-OCT-2001; 2001US-0327449P.
XX PR 09-OCT-2001; 2001US-0327917P.
XX PR 09-OCT-2001; 2001US-0328029P.
XX PR 09-OCT-2001; 2001US-0328044P.
XX PR 09-OCT-2001; 2001US-0328056P.
XX PR 12-OCT-2001; 2001US-0328849P.

PR 15-OCT-2001; 2001US-0329414P.
PR 17-OCT-2001; 2001US-0330142P.
PR 18-OCT-2001; 2001US-0330309P.
PR 22-OCT-2001; 2001US-0341058P.
PR 24-OCT-2001; 2001US-0339266P.
PR 24-OCT-2001; 2001US-0343629P.
PR 29-OCT-2001; 2001US-0349575P.
PR 01-NOV-2001; 2001US-0346357P.
PR 17-APR-2002; 2002US-0373260P.
PR 19-APR-2002; 2002US-0373815P.
PR 19-APR-2002; 2002US-0373817P.
PR 19-APR-2002; 2002US-0373826P.
PR 19-APR-2002; 2002US-0373894P.
PR 22-APR-2002; 2002US-0374977P.
PR 16-MAY-2002; 2002US-0381037P.
PR 16-MAY-2002; 2002US-0381038P.
PR 16-MAY-2002; 2002US-0381042P.
PR 17-MAY-2002; 2002US-0381642P.
PR 28-MAY-2002; 2002US-0383656P.
PR 29-MAY-2002; 2002US-0383831P.
PR 25-JUN-2002; 2002US-0391335P.
PR 01-OCT-2002; 2002US-00262511.
XX
XX (CURA-) CURAGEN CORP.
XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;
PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dippio VA;
PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;
XX WPI; 2003-381626/36.
XX N-PSDB; ADA05741.
DR
XX New NOVX polypeptides and nucleic acids, useful for diagnosing,
PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
XX Claim 1; Page 172; 586pp; English.
XX
CC The present invention describes NOVX proteins, where X can be 1 to 55
CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide
CC described above and a carrier; (2) a kit comprising, in one or more
CC containers, the composition described above; (3) an isolated nucleic acid
CC molecule which encodes a NOVX protein of the invention; (4) a vector
CC comprising the nucleic acid molecule described above; (5) a cell
CC comprising the above vector; (6) an antibody that immunospecifically
CC binds to the polypeptide described above; (7) methods for determining the
CC presence or amount of the above polypeptide or nucleic acid molecule in a
CC sample; (8) methods for determining the presence of or predisposition to
CC a disease associated with altered levels of expression of the above
CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
CC method of identifying an agent that binds to the polypeptide described
CC above; (10) a method for identifying a potential therapeutic agent for
CC use in treating a pathology that is related to an aberrant expression or
CC aberrant physiological interactions of the polypeptide; (11) a method of
CC screening for a modulator of activity or of latency or predisposition to
CC a pathology associated with the polypeptide; (12) a method for modulating
CC the activity of the polypeptide described above; (13) methods of treating
CC or preventing a pathology associated with the above polypeptide in a
CC mammal; and (14) a method for producing the above polypeptide. NOVX
CC sequences have antidiabetic, anorectic, antibacterial, virucide,
CC immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian
CC and antilipemic activities, and can be used in gene therapy. The
CC polypeptide is useful in manufacturing a medicament for treating a
CC syndrome associated with a human disease. The polypeptide or the nucleic
CC acid molecule may be used to diagnose, treat or prevent metabolic
CC disorders such as diabetes or obesity, infections, cachexia, cancer,
CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
CC disease, immune disorders, haematopoietic disorders and various
CC dyslipidaemias. The nucleic acids can also be used as hybridisation

CC probes, in chromosome mapping, tissue typing, preventive medicine and
 CC pharmacogenomics. The present sequence represents a human NOVX from the
 CC present invention.

XX Sequence 247 AA;

SQ Query Match 100.0%; Score 40; DB 6; Length 247;
 Best Local Similarity 100.0%; Pred. No. 2.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMWKKV 9
 Db 109 RLSSMWKKV 117

RESULT 5
 ADA05732
 ID ADA05732 standard; protein; 250 AA.

AC ADA05732;

DT 06-NOV-2003 (first entry)

DE Human NOV18a protein SEQ ID NO:92.

XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
 KW immunomodulator; cytostatic; neurotropic; neuroprotective;
 KW antiparkinsonian; antilipemic; gene therapy; human disease;
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; dyslipidaemia.

XX Homo sapiens.

XX WO2003029424-A2.

XX 10-APR-2003.

XX 02-OCT-2002; 2002WO-US031373.

XX 02-OCT-2001; 2001US-0326483P.

XX 05-OCT-2001; 2001US-0327435P.

XX 05-OCT-2001; 2001US-0327449P.

XX 09-OCT-2001; 2001US-0327917P.

XX 09-OCT-2001; 2001US-0328029P.

XX 09-OCT-2001; 2001US-0328044P.

XX 09-OCT-2001; 2001US-0328056P.

XX 12-OCT-2001; 2001US-0328849P.

XX 15-OCT-2001; 2001US-0329414P.

XX 17-OCT-2001; 2001US-0330142P.

XX 18-OCT-2001; 2001US-0330309P.

XX 22-OCT-2001; 2001US-0341058P.

XX 24-OCT-2001; 2001US-0339266P.

XX 24-OCT-2001; 2001US-0343629P.

XX 29-OCT-2001; 2001US-0349575P.

XX 01-NOV-2001; 2001US-0346357P.

XX 17-APR-2002; 2002US-0373260P.

XX 19-APR-2002; 2002US-0373815P.

XX 19-APR-2002; 2002US-0373817P.

XX 19-APR-2002; 2002US-0373826P.

XX 19-APR-2002; 2002US-0373884P.

XX 22-APR-2002; 2002US-0374977P.

XX 16-MAY-2002; 2002US-0381037P.

XX 16-MAY-2002; 2002US-0381038P.

XX 16-MAY-2002; 2002US-0381042P.

XX 17-MAY-2002; 2002US-0381642P.

XX 28-MAY-2002; 2002US-0383656P.

XX 29-MAY-2002; 2002US-0383831P.

XX 25-JUN-2002; 2002US-0391335P.

XX 01-OCT-2002; 2002US-00262511.

XX (CURA-) CURAGEN CORP.

XX

PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
 PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
 PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;
 PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;
 PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;
 XX WPI: 2003-381526/36.
 DR N-PSDB; ADA05731.
 XX New NOVX polypeptides and nucleic acids, useful for diagnosing,
 PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
 PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.
 XX Claim 1; Page 169-170; 586pp; English.

PS The present invention describes NOVX proteins, where X can be 1 to 55
 XX (e.g. NOV1). Also described: (1) a composition comprising a polypeptide
 CC described above and a carrier; (2) a kit comprising, in one or more
 CC containers, the composition described above; (3) an isolated nucleic acid
 CC molecule which encodes a NOVX protein of the invention; (4) a vector
 CC comprising the nucleic acid molecule described above; (5) a cell
 CC binds to the polypeptide described above; (6) an antibody that immunospecifically
 CC binds to the polypeptide described above; (7) methods for determining the
 CC presence or amount of the above polypeptide or nucleic acid molecule in a
 CC sample; (8) methods for determining the presence of or predisposition to
 CC a disease associated with altered levels of expression of the above
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
 CC method of identifying an agent that binds to the polypeptide described
 CC above; (10) a method for identifying a potential therapeutic agent for
 CC use in treating a pathology that is related to an aberrant expression or
 CC aberrant physiological interactions of the polypeptide; (11) a method of
 CC screening for a modulator of activity or of latency or predisposition to
 CC a pathology associated with the polypeptide; (12) a method for modulating
 CC the activity of the polypeptide described above; (13) methods of treating
 CC or preventing a pathology associated with the above polypeptide in a
 CC mammal; and (14) a method for producing the above polypeptide. NOVX
 CC sequences have antidiabetic, anorectic, neurotropic, antiparkinsonian
 CC immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian
 CC and antilipemic activities, and can be used in gene therapy. The
 CC polypeptide is useful in manufacturing a medicament for treating a
 CC syndrome associated with a human disease. The polypeptide or the nucleic
 CC acid molecule may be used to diagnose, treat or prevent metabolic
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
 CC disease, immune disorders, haematopoietic disorders and various
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation
 CC probes, in chromosome mapping, tissue typing, preventive medicine and
 CC pharmacogenomics. The present sequence represents a human NOVX from the
 CC present invention.

XX Sequence 250 AA;

SQ Query Match 100.0%; Score 40; DB 6; Length 250;

Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RLSSMWKKV 9

Db 120 RLSSMWKKV 128

RESULT 6

ADA05734

ID ADA05734 standard; protein; 252 AA.

XX ADA05734;

XX 06-NOV-2003 (first entry)

XX Human NOV18b protein SEQ ID NO:94.

XX

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:16:55 ; Search time 45.5556 Seconds
(without alignments)
55.820 Million cell updates/sec

Title: US-09-905-083-86
Perfect score: 41
Sequence: 1 MARSLPL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	9	4 AAE08291	Human str
2	41	100.0	136	4 ABG23378	Novel hum
3	41	100.0	198	6 ADA05736	Human NOV
4	41	100.0	253	2 AAR67888	Human str
5	41	100.0	253	2 AAW05383	Human any
6	41	100.0	253	5 ABB84421	Human SCC
7	41	100.0	253	5 ABB84406	Human SCC
8	41	100.0	253	5 AAU82740	Amino aci
9	41	100.0	253	6 ABU07440	Protein d
10	41	100.0	253	6 ABU07471	Human str
11	41	100.0	253	6 AB58471	Human str
12	41	100.0	253	7 ADB80484	Ovarian c
13	41	100.0	257	3 AAB21326	Human HSC
14	36	87.8	9	4 AAE08320	Human str
15	36	87.8	818	5 ABB98135	Human PMM
16	36	87.8	18	5 AAU82753	Amino aci
17	35	85.4	475	4 AAE10657	Secreted
18	35	85.4	475	4 AAE02609	Human sec
19	35	85.4	475	5 ABB78618	Secreted
20	35	85.4	518	2 AAW61362	Aspartic
21	35	85.4	518	2 AAY13799	Human str
22	35	85.4	518	2 AAY22239	Human CSP
23	35	85.4	518	2 AAY41714	Human PRO
24	35	85.4	518	3 AAY88424	Human asp
25	35	85.4	518	3 AAB44270	Human PRO

26	35	85.4	518	4 AAU07201	Human asp
27	35	85.4	518	4 AAE10628	Human asp
28	35	85.4	518	4 AAE10656	Human-Asp
29	35	85.4	518	4 AAE06858	Human asp
30	35	85.4	518	4 AAE02608	Human asp
31	35	85.4	518	4 AAE02580	Human asp
32	35	85.4	518	4 AAU29059	Human PRO
33	35	85.4	518	4 AAU06602	Human Asp
34	35	85.4	518	5 ABB06531	Human asp
35	35	85.4	518	5 ABB78589	Human Asp
36	35	85.4	518	5 ABB78617	Human BAC
37	35	85.4	518	5 ABB07453	Human Asp
38	35	85.4	518	6 ABB58435	Human PRO
39	35	85.4	518	6 ABU87983	Novel hum
40	35	85.4	518	6 ABU84298	Human sec
41	35	85.4	518	6 ABR66172	Human sec
42	35	85.4	518	6 ABR65562	Human sec
43	35	85.4	518	6 ABU99502	Human sec
44	35	85.4	518	6 ABU82741	Human PRO
45	35	85.4	518	6 ABU89862	Novel hum

ALIGNMENTS

RESULT 1

AAE08291 standard; peptide; 9 AA.
XX AAE08291;
XX AAE08291;
XX 01-NOV-2001 (first entry)
XX Human stratum corneum chymotrypsin enzyme peptide #56 (residues 1-9).
XX Stratum corneum chymotrypsin enzyme; SCCE; cytostatic; vaccine; tumour;
XX cancer; ovarian; breast; lung; colon; prostate; carcinoma; human;
XX antiseptic therapy; malignant hyperplasia.
XX Homo sapiens.
XX WO200159158-A1.
XX 16-AUG-2001.
XX 07-FEB-2001; 2001WO-US003977.
XX 11-FEB-2000; 2000US-00502600.
XX (UYAR-) UNIV ARKANSAS.
XX O'Brien TJ;
XX WPI; 2001-514676/56.
XX Diagnosing cancer comprises detecting stratum corneum chymotrypsin enzyme.
XX Claim 25; Page 114; 127pp; English.
XX The invention relates to diagnosing cancer especially ovarian cancer, by screening for stratum corneum chymotrypsin enzyme (SCCE). Proteases are considered to be an integral part of their tumour growth and metastasis, and therefore, markers indicative of their presence or absence are useful for the diagnosis of cancer. The method is useful for diagnosing cancer, CC carcinoma and malignant hyperplasia. The SCCE oligonucleotide may be used CC to treat a cancer selected from ovarian, breast, lung, colon, prostate CC and other cancers in which SCCE is overexpressed. The present sequence is CC human SCCE peptide

XX Sequence 9 AA;

Query Match 100.0%; Score 41; DB 4; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLEPL 9
| | | | |
Db 1 MARSLLEPL 9

RESULT 2
ABG23378
ID ABG23378 standard; protein; 136 AA.

XX AC ABG23378;
XX 18-FEB-2002 (first entry)
XX Novel human diagnostic protein #23369.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.

XX WC200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSB-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS87565.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX Claim 20; SEQ ID NO 53737; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG0377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 136 AA;

XX Query Match 100.0%; Score 41; DB 4; Length 136;
XX Best Local Similarity 100.0%; Pred. No. 1.5;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLEPL 9
| | | | |
Db 7 MARSLLEPL 15

RESULT 3
ADA05736
ID ADA05736 standard; protein; 198 AA.

XX AC ADA05736;

XX 06-NOV-2003 (first entry)

XX Human NOV18c protein SEQ ID NO:96.

XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
XX immunomodulator; cytostatic; neurotropic; neuroprotective;
XX antiparkinsonian; antilipemic; gene therapy; human disease;
XX metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX immune disorder; haematopoietic disorder; dyslipidaemia.

XX OS Homo sapiens.

XX WO2003029424-A2.

XX 10-APR-2003.

XX 02-OCT-2002; 2002WO-US031373.

XX 02-OCT-2001; 2001US-0326483P.

XX 05-OCT-2001; 2001US-0327435P.

XX 09-OCT-2001; 2001US-0327449P.

XX 09-OCT-2001; 2001US-0327917P.

XX 09-OCT-2001; 2001US-0328029P.

XX 09-OCT-2001; 2001US-0328044P.

XX 12-OCT-2001; 2001US-0328056P.

XX 15-OCT-2001; 2001US-0328849P.

XX 17-OCT-2001; 2001US-0329414P.

XX 18-OCT-2001; 2001US-0330142P.

XX 22-OCT-2001; 2001US-0330309P.

XX 24-OCT-2001; 2001US-0341058P.

XX 24-OCT-2001; 2001US-0339286P.

XX 29-OCT-2001; 2001US-0343629P.

XX 01-NOV-2001; 2001US-0349575P.

XX 17-APR-2002; 2002US-0373260P.

XX 19-APR-2002; 2002US-0373815P.

XX 19-APR-2002; 2002US-0373817P.

XX 19-APR-2002; 2002US-0373826P.

XX 22-APR-2002; 2002US-0373884P.

XX 16-MAY-2002; 2002US-0374977P.

XX 16-MAY-2002; 2002US-0381037P.

XX 16-MAY-2002; 2002US-0381038P.

XX 17-MAY-2002; 2002US-0381042P.

XX 28-MAY-2002; 2002US-0381642P.

XX 29-MAY-2002; 2002US-0383656P.

XX 25-JUN-2002; 2002US-0383831P.

XX 01-OCT-2002; 2002US-0391335P.

XX (CURA-) CURAGEN CORP.

XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;

XX Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;

XX Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;

XX Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;

XX Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Bergus C, Dipippo VA;

XX Eisen AJ, Gargolli EA, Kieger DK, Spaderna SK;

XX WPI; 2003-381626/36.

XX N-PSDB; ADA05735.

XX New NOVX polypeptides and nucleic acids, useful for diagnosing,
PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
XX Claim 1; Page 170; 586pp; English.
XX
XX The present invention describes NOVX proteins, where X can be 1 to 55
CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide
CC described above and a carrier; (2) a kit comprising, in one or more
CC containers, the composition described above; (3) an isolated nucleic acid
CC molecule which encodes a NOVX protein of the invention; (4) a vector
CC comprising the nucleic acid molecule described above; (5) a cell
CC comprising the above vector; (6) an antibody that immunospecifically
CC binds to the polypeptide described above; (7) methods for determining the
CC presence or amount of the above polypeptide or nucleic acid molecule in a
CC sample; (8) methods for determining the presence of or predisposition to
CC a disease associated with altered levels of expression of the above
CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
CC method of identifying an agent that binds to the polypeptide described
CC above; (10) a method for identifying a potential therapeutic agent for
CC use in treating a pathology that is related to an aberrant expression or
CC aberrant physiological interactions of the polypeptide; (11) a method of
CC screening for a modulator of activity of or of latency or predisposition to
CC a pathology associated with the polypeptide; (12) a method for modulating
CC the activity of the polypeptide described above; (13) methods of treating
CC or preventing a pathology associated with the above polypeptide in a
CC mammal; and (14) a method for producing the above polypeptide. NOVX
CC sequences have antidiabetic, anorectic, antibacterial, virucide,
CC immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian
CC and antipapillary activities, and can be used in gene therapy. The
CC polypeptide is useful in manufacturing a medicament for treating a
CC syndrome associated with a human disease. The polypeptide or the nucleic
CC acid molecule may be used to diagnose, treat or prevent metabolic
CC disorders such as diabetes or obesity, infections, cachexia, cancer,
CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
CC disease, immune disorders, haematopoietic disorders and various
CC dyslipidaemias. The nucleic acids can also be used as hybridisation
CC probes, in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. The present sequence represents a human NOVX from the
XX present invention.
XX
XX Sequence 198 AA;
SQ
Query Match 100.0%; Score 41; DB 6; Length 198;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MARSLLLPL 9
DB 1 MARSLLLPL 9
RESULT 4
AAR67888
ID AAR67888 standard; protein; 253 AA.
XX
XX AAR67888;
AC
XX 25-MAR-2003 (revised)
DT 09-AUG-1995 (first entry)
XX
XX Human stratum corneum chymotryptic recombinant enzyme (SCCE).
XX Stratum corneum chymotryptic enzyme; skin disorder; acne; psoriasis;
XX callosities; keratosis pilaris; ichthyoses; eczema.
XX
XX Homo sapiens.
OS
XX WO9500651-A1.
PN
XX 05-JAN-1995.
PD

XX 20-JUN-1994; 94WO-IB000166.
XX
XX 18-JUN-1993; 93DK-0000725.
XX
XX (SYMB-) SYMBICOM AB.
PA
XX Egeirud T, Hansson L;
PI
XX WPI; 1995-052088/07.
DR
XX N-PSDB; AAQ81203.
XX
XX Nucleotide sequences encoding stratum corneum chymotryptic enzyme - and
PT related vectors, transformed cells and polypeptides, useful for treating
PT skin disorders, e.g. acne or psoriasis, and for identification of
PT specific inhibitors.
XX
XX Disclosure; Page 97; 137pp; English.
XX
XX The enzyme encoded by this sequence is used in pharmaceutical, cosmetic
CC and skin care products, especially to treat and prevent acne, xeroderma,
CC or other hyperkeratotic conditions (e.g. callosities or keratosis
CC pilaris), ichthyoses, psoriasis, eczema, etc. It is produced
CC recombinantly following mammal, insect, plant, or microorganism
CC transformation with plasmid pS507. (Updated on 25-MAR-2003 to correct PN
XX field.)
XX
XX Sequence 253 AA;
SQ
Query Match 100.0%; Score 41; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MARSLLLPL 9
DB 1 MARSLLLPL 9
RESULT 5
AAW05383
ID AAW05383 standard; protein; 253 AA.
XX
XX AAW05383;
AC
XX 31-DEC-1996 (first entry)
DT
XX Human amyloid precursor protein protease.
XX
XX Amyloid precursor protein protease; Alzheimer's disease; diagnosis;
XX therapy.
XX
XX Homo sapiens.
OS
XX WO9631122-A1.
PN
XX 10-OCT-1996.
PD
XX 02-APR-1996; 96WO-US004294.
PF
XX 04-APR-1995; 95US-00416257.
PR
XX (ELIL) LILLY & CO ELI.
PA
XX Dixon EP, Johnstone EM, Little SP;
PI
XX WPI; 1996-464694/46.
DR
XX N-PSDB; AAT39783.
XX
XX New isolated human amyloid precursor protein protease - used to develop
PT prods. for the treatment or diagnosis of associated conditions, esp.
PT Alzheimer's disease.
XX
XX Claim 1; Page 44-45; 55pp; English.
PS

XX Human amyloid precursor protein protease (AAW05383) is involved in the
 CC processing or clearance of amyloid precursor protein to form beta-amyloid
 CC peptide. Its amino acid sequence was deduced from a cDNA clone (AA39783)
 CC ctd. from a human lung library. Recombinant protease can be produced in
 CC transformed or transfected prokaryotic (partic. E. coli) or eukaryotic
 CC (partic. AV-120 host cells. It is used to develop products for the design
 CC and testing of cpds. useful for treating or preventing conditions
 CC associated with beta-amyloid peptide, esp. Alzheimer's disease
 XX
 SQ Sequence 253 AA;

Query Match 100.0%; Score 41; DB 2; Length 253;
 Best Local Similarity 100.0%; Pred. No. 2.9;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
 |||||
 Db 1 MARSLLLPL 9

RESULT 6
 ABB84421
 ID ABB84421 standard; peptide; 253 AA.

XX ABB84421;

DT 08-NOV-2002 (first entry)

DE Human SCCE protein N-terminal fragment SEQ ID 48.

XX SCCE; human; stratum corneum chymotryptic enzyme; kallikrein 7;
 KW serine protease; transgenic mammal; skin; skin disease; skin cancer;
 KW hyperkeratosis; acanthosis; epidermal inflammation; dermal inflammation;
 KW pruritus; atopic dermatitis; eczema; acne; itch; KLK7; enzyme.

XX Homo sapiens.

OS WO200262135-A2.

PN 15-AUG-2002.

PD 08-FEB-2002; 2002WO-IB001300.

PF 09-FEB-2001; 2001CA-02332655.

PR 09-FEB-2001; 2001DK-00000218.

XX (EGEL/) EGELRUD T.

PA (HANS/) HANSSON L.

XX Egelrud T, Hansson L;

XX WPI; 2002-643380/59.

XX Transgenic mammal or its embryo useful as model for human disease, has
 PT heterologous nucleotide sequence coding for stratum corneum chymotryptic
 PT enzyme operably linked to promoter that drives its expression in skin.

XX Example 6; Page 37; 74pp; English.

XX This invention describes a novel non-human transgenic mammal or mammalian
 CC embryo having integrated within its genome, a heterologous nucleotide
 CC sequence comprising at least a significant part of a nucleotide sequence
 CC coding for a stratum corneum chymotryptic enzyme (SCCE) or its variant,
 CC operably linked to a promoter that drives expression of heterologous scce
 CC or its variant in skin. The product of the invention is useful as a model
 CC for the study of disease with the aim of improving treatment, to relieve
 CC or ameliorate a pathogenic condition, for development or testing of a
 CC cosmetic or a pharmaceutical formulation, and for the development of a
 CC diagnostic method. It can also be used as a model for a skin disease or
 CC skin cancer. The invention is also useful for screening or identifying a
 CC compound or composition effective for the prevention or treatment of an
 CC abnormal or unwanted phenotype, and for screening or identifying a

CC compound or composition effective for the prevention or treatment of
 CC inflammatory skin diseases selected from diseases consisting of epidermal
 CC hyperkeratosis, acanthosis, epidermal inflammation, dermal inflammation,
 CC pruritus, atopic dermatitis, eczema, acne and inherited skin diseases
 CC with epidermal hyperkeratosis. The mammal of the invention is also useful
 CC as a model for further studies of its mechanisms and the testing of
 CC potential compounds and compositions for relieve of various skin diseases
 CC where itch is a component. This sequence represents the N-terminal
 CC fragment of the human stratum corneum chymotryptic enzyme, SCCE
 CC synonymous with human kallikrein 7 (KLK7) used in the development of the
 CC transgenic mammals described in the invention
 XX

SQ Sequence 253 AA;

Query Match 100.0%; Score 41; DB 5; Length 253;

Best Local Similarity 100.0%; Pred. No. 2.9;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
 |||||
 Db 1 MARSLLLPL 9

RESULT 7

ABB84406

ID ABB84406 standard; protein; 253 AA.

XX ABB84406;

DT 08-NOV-2002 (first entry)

DE Human SCCE protein.

XX SCCE; human; stratum corneum chymotryptic enzyme; kallikrein 7;

KW serine protease; transgenic mammal; skin; skin disease; skin cancer;

KW hyperkeratosis; acanthosis; epidermal inflammation; dermal inflammation;

KW pruritus; atopic dermatitis; eczema; acne; itch; KLK7; enzyme.

XX Homo sapiens.

OS WO200262135-A2.

PN 15-AUG-2002.

PD 08-FEB-2002; 2002WO-IB001300.

PF 09-FEB-2001; 2001CA-02332655.

PR 09-FEB-2001; 2001DK-00000218.

XX (EGEL/) EGELRUD T.

PA (HANS/) HANSSON L.

XX Egelrud T, Hansson L;

XX WPI; 2002-643380/59.

XX N-PSDB; ABQ76226.

XX Transgenic mammal or its embryo useful as model for human disease, has
 PT heterologous nucleotide sequence coding for stratum corneum chymotryptic
 PT enzyme operably linked to promoter that drives its expression in skin.
 XX Claim 10; Page 58-59; 74pp; English.
 XX This invention describes a novel non-human transgenic mammal or mammalian
 CC embryo having integrated within its genome, a heterologous nucleotide
 CC sequence comprising at least a significant part of a nucleotide sequence
 CC coding for a stratum corneum chymotryptic enzyme (SCCE) or its variant,
 CC operably linked to a promoter that drives expression of heterologous scce
 CC or its variant in skin. The product of the invention is useful as a model
 CC for the study of disease with the aim of improving treatment, to relieve
 CC or ameliorate a pathogenic condition, for development or testing of a
 CC cosmetic or a pharmaceutical formulation, and for the development of a
 CC diagnostic method. It can also be used as a model for a skin disease or

CC skin cancer. The invention is also useful for screening or identifying a
 CC compound or composition effective for the prevention or treatment of an
 CC abnormal or unwanted phenotype, and for screening or identifying a
 CC compound or composition effective for the prevention or treatment of
 CC inflammatory skin diseases selected from diseases consisting of epidermal
 CC hyperkeratosis, acanthosis, epidermal inflammation, dermal inflammation,
 CC pruritus, atopic dermatitis, eczema, acne and inherited skin diseases.
 CC with epidermal hyperkeratosis. The mammal of the invention is also useful
 CC as a model for further studies of itch mechanisms and the testing of
 CC potential compounds and compositions for relieve of various skin diseases
 CC where itch is a component. This sequence represents the human stratum
 CC corneum chymotryptic enzyme, SCCE which is a serine protease synonymum
 CC with human kallikrein 7 (KLK7) and is used in the development of the
 CC transgenic mammals described in the invention
 XX
 XX Sequence 253 AA;

Query Match 100.0%; Score 41; DB 5; Length 253;
 Best Local Similarity 100.0%; Pred. No. 2.9;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9
 |||||
 Db 1 MARSLLLPL 9

RESULT 8
 AAU82740
 ID AAU82740 standard; protein; 253 AA.

XX AC AAU82740;

XX DT 23-APR-2002 (first entry)

XX DE Amino acid sequence of novel human protease #39.

XX Human; protease; cancer; immune-related disorder; cardiovascular disease;
 KW neuronal-associated disease; metabolic disorder; inflammatory disorder;
 KW nervous system disorder; sexual dysfunction; pain; mood disorder;
 KW hyperension; psychotic disorder; neurological disorder; dyskinesia;
 KW viral infection; human immunodeficiency virus; HIV; non-viral infection;
 KW ocular disease; cytostatic; enzyme.

XX OS Homo sapiens.

XX PN WO200200860-A2.

XX PD 03-JAN-2002.

XX PF 26-JUN-2001; 2001WO-US020171.

XX PR 26-JUN-2000; 2000US-0214047P.

XX PA (SUGEN-) SUGEN INC.

XX PI Plowman G, Whyte D, Sudarsanam S, Manning G, Caenepeel S;

XX PI Charyczak G;

XX DR WPI; 2002-139913/18.

XX DR N-PSDB; ABK31782.

XX Nucleic acids encoding novel human proteases, useful for useful for
 PT treating diseases and disorders such as cancers, immune-related diseases
 PT and disorders, cardiovascular disease (e.g. restenosis) and inflammatory
 PT disorders.

XX PS Claim 6; Fig 2N; 313pp; English.

XX The present invention relates to the isolation of novel human proteases,
 CC and the nucleic acids encoding them. The sequences of the invention are
 CC useful for treating diseases and disorders such as cancers (e.g. breast,
 CC colon, lung), immune-related diseases and disorders (e.g. inflammatory
 CC diseases and asthma), cardiovascular diseases (e.g. restenosis and

CC coronary thrombosis), brain or neuronal-associated diseases, metabolic
 CC disorders (e.g. diabetes, obesity), inflammatory disorders (e.g.
 CC rheumatoid arthritis and psoriasis), central or peripheral nervous system
 CC diseases, migraines, pain, sexual dysfunction, mood disorders, attention
 CC disorders, cognition disorders, hypotension, hypertension, psychotic
 CC diseases, neurological disorders (e.g. Alzheimer's disease, Parkinson's
 CC disease) and dyskinesias. The nucleic acids and polypeptides are also
 CC useful for treating viral infections caused by human immunodeficiency
 CC virus (HIV), and non-viral infections such as ocular disease (e.g.
 CC glaucoma) and macular degeneration. AAU82702-AAU82760 represent the novel
 CC human proteases of the invention
 XX

XX Sequence 253 AA;

Query Match 100.0%; Score 41; DB 5; Length 253;
 Best Local Similarity 100.0%; Pred. No. 2.9;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9
 |||||
 Db 1 MARSLLLPL 9

RESULT 9
 ABU07440

ID ABU07440 standard; protein; 253 AA.

XX AC ABU07440;

XX DT 28-JAN-2003 (first entry)

XX DE Protein differentially regulated in prostate cancer #43.

XX Prostate cancer; gene expression; differential regulation;
 KW molecular marker; drug target; cancer detection; cancer diagnosis;
 KW cancer staging; cancer grading; cancer assessing; cancer monitoring.

XX OS Homo sapiens.

XX PN WO200281638-A2.

XX PD 17-OCT-2002.

XX PF 08-APR-2002; 2002WO-US010824.

XX PR 06-APR-2001; 2001US-0281731P.

XX PR 06-APR-2001; 2001US-0281732P.

XX PA (ORIG-) ORIGENE TECHNOLOGIES INC.

XX PI Sun Z, Jay G;

XX DR WPI; 2003-058520/05.

XX DR N-PSDB; ABX10343.

XX Novel genes which are differentially regulated in prostate cancer, useful
 PT for diagnosing prostate cancer in prostate tissue sample and assessing
 PT therapeutic or preventive intervention in prostate cancer patients.

XX Claim 1; Page 293-294; 416pp; English.

XX The invention describes genes (I) which are differentially regulated in
 CC prostate cancer. (I) is useful for diagnosing a prostate cancer in a
 CC sample comprising prostate tissue, which involves determining the number
 CC of target genes which are differentially-regulated in the sample, where
 CC the number is indicative of the probability that the sample comprises
 CC prostate cancer. (I) is useful for assessing a therapeutic or preventive
 CC intervention in a subject having a prostate cancer, which involves
 CC determining the expression levels in a sample comprising prostate tissue
 CC of target genes which are differentially-regulated in prostate cancer.
 CC Preferably, the expression levels of at least 10 genes are determined.
 CC (I) is also useful for identifying agents that modulate a biological
 CC activity of a polypeptide differentially-regulated in prostate cancer

CC cells, which involves contacting a polypeptide differentially-regulated
 CC in prostate cancer cells with a test agent under conditions effective for
 CC the test agent to modulate a biological activity of the polypeptide, and
 CC determining whether the test agent modulates the biological activity. (I)
 CC is useful as molecular markers, as drug targets, and for detecting,
 CC diagnosing, staging, grading, assessing, monitoring, prognosticating,
 CC preventing or treating, determining predisposition to diseases and
 CC conditions especially relating to prostate cancer. (I) and its expression
 CC products are used in the diagnostic test to assay for presence of cancer
 CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
 CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
 CC of cancer, its stage of development, the nature of genetic defect, etc.
 CC The polypeptide encoded by (I) can be used as target for therapy or drug
 CC discovery. (I) can also be used for expressing the polypeptide and thus
 CC for searching specific binding partners of the polypeptide. (I) is useful
 CC in therapeutic applications to treat prostate cancer. The identification
 CC of specific genes, and groups of genes, expressed in pathways
 CC physiologically relevant to prostate cancer permits the definition of
 CC functional and disease pathways and the delineation of targets in these
 CC pathways which are useful in diagnostic, therapeutic, and clinical
 CC applications. This is the amino acid sequence of a protein differentially
 CC regulated in prostate cancer
 XX
 SQ Sequence 253 AA;

Query Match 100.0%; Score 41; DB 6; Length 253;
 Best Local Similarity 100.0%; Pred. No. 2.9;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9
 |||||
 Db 1 MARSLLLPL 9

RESULT 10

ABU07471
 ID ABU07471 standard; protein; 253 AA.

AC ABU07471;

DT 28-JAN-2003 (first entry)

DE Protein differentially regulated in prostate cancer #74.

XX Prostate cancer; gene expression; differential regulation;
 KW molecular marker; drug target; cancer detection; cancer diagnosis;
 KW cancer staging; cancer grading; cancer assessing; cancer monitoring.

XX Homo sapiens.

XX WO200281638-A2.

XX 17-OCT-2002.

XX 08-APR-2002; 2002WO-US010824.

XX 06-APR-2001; 2001US-0281731P.

XX 06-APR-2001; 2001US-0281732P.

XX (ORIG-) ORIGENE TECHNOLOGIES INC.

XX Sun Z, Jay G;

XX WPI; 2003-058520/05.

XX N-PSDB; ABX10375.

XX Novel genes which are differentially regulated in prostate cancer, useful
 PT for diagnosing prostate cancer in prostate tissue sample and assessing
 PT therapeutic or preventive intervention in prostate cancer patients.

XX Claim 1; Page 351; 416pp; English.

XX The invention describes genes (I) which are differentially regulated in

CC prostate cancer. (I) is useful for diagnosing a prostate cancer in a
 CC sample comprising prostate tissue, which involves determining the number
 CC of target genes which are differentially-regulated in the sample, where
 CC the number is indicative of the probability that the sample comprises
 CC prostate cancer. (I) is useful for assessing a therapeutic or preventive
 CC intervention in a subject having a prostate cancer, which involves
 CC determining the expression levels in a sample comprising prostate tissue
 CC of target genes which are differentially-regulated in prostate cancer.
 CC Preferably, the expression levels of at least 10 genes are determined.
 CC (I) is also useful for identifying agents that modulate a biological
 CC activity of a polypeptide differentially-regulated in prostate cancer
 CC cells, which involves contacting a polypeptide differentially-regulated
 CC in prostate cancer cells with a test agent under conditions effective for
 CC the test agent to modulate a biological activity of the polypeptide, and
 CC determining whether the test agent modulates the biological activity. (I)
 CC is useful as molecular markers, as drug targets, and for detecting,
 CC diagnosing, staging, grading, assessing, monitoring, prognosticating,
 CC preventing or treating, determining predisposition to diseases and
 CC conditions especially relating to prostate cancer. (I) and its expression
 CC products are used in the diagnostic test to assay for presence of cancer
 CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
 CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
 CC of cancer, its stage of development, the nature of genetic defect, etc.
 CC The polypeptide encoded by (I) can be used as target for therapy or drug
 CC discovery. (I) can also be used for expressing the polypeptide and thus
 CC for searching specific binding partners of the polypeptide. (I) is useful
 CC in therapeutic applications to treat prostate cancer. The identification
 CC of specific genes, and groups of genes, expressed in pathways
 CC physiologically relevant to prostate cancer permits the definition of
 CC functional and disease pathways and the delineation of targets in these
 CC pathways which are useful in diagnostic, therapeutic, and clinical
 CC applications. This is the amino acid sequence of a protein differentially
 CC regulated in prostate cancer
 XX
 SQ Sequence 253 AA;

Query Match 100.0%; Score 41; DB 6; Length 253;

Best Local Similarity 100.0%; Pred. No. 2.9;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9

Db 1 MARSLLLPL 9

RESULT 11

ABR58471

ID ABR58471 standard; protein; 253 AA.

XX ABR58471;

XX 07-JUL-2003 (first entry)

XX Human stratum corneum chymotryptic enzyme - ovarian cancer clone O1676P.

XX Human; cytostatic; gene therapy; vaccine; cancer; ovarian cancer.

XX Homo sapiens.

XX WO2003029468-A1.

XX 10-APR-2003.

XX 02-OCT-2002; 2002WO-US031467.

XX 02-OCT-2001; 2001US-0327135P.

XX 30-MAY-2002; 2002US-0384531P.

XX (CORI-) CORIXA CORP.

XX Algate PA, Mannion J;

XX WPI; 2003-372001/35.

XX DR

XX New polynucleotide and polypeptide useful for diagnosing and/or treating
PT cancer, particularly ovarian cancer, and as a vaccine.
XX
XX
PS Claim 2; Page 157-158; 169pp; English.

XX The invention relates to a novel isolated polynucleotide. The
CC polynucleotides of the invention have cytostatic activity, and may have a
CC use in gene therapy, and in a vaccine. The composition and methods are
CC useful in diagnosing and/or treating cancer, particularly ovarian cancer.
CC The composition may also be used as a vaccine to prevent cancer. The
CC present sequence is used in the exemplification of the invention
XX
XX Sequence 253 AA;

Query Match 100.0%; Score 41; DB 6; Length 253;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
Db 1 MARSLLLPL 9

RESULT 12
ADB80484
ID ADB80484 standard; protein; 253 AA.

XX AC ADB80484;
XX DT 04-DEC-2003 (first entry)
XX DE Ovarian cancer-associated protein #24.

XX cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
XX post-operative chemotherapy; radiation therapy; tumour prognosis;
XX pre-cancerous lesion detection.
XX Homo sapiens.
XX WO2002102235-A2.
XX PD 27-DEC-2002.
XX PF 18-JUN-2002; 2002WO-US019297.
XX PR 18-JUN-2001; 2001US-0299234P.
XX PR 27-AUG-2001; 2001US-0315287P.
XX PR 05-SEP-2001; 2001US-0317544P.
XX PR 13-NOV-2001; 2001US-0350666P.
XX PR 12-APR-2002; 2002US-0372246P.

XX PA (BOSB-) EOS BIOTECHNOLOGY INC.
XX Mack DH, Gish KC;
XX WI; 2003-167431/16.
XX N-PSDB; ADB80483.
XX Detecting an ovarian cancer-associated transcript in a cell from a
PT patient, comprises contacting a biological sample from the patient with a
PT polynucleotide that hybridizes to an ovarian cancer gene.
XX
XX Claim 13; Page 291; 332pp; English.

XX The invention relates to a method of detecting an ovarian cancer-
CC associated transcript in a cell from a patient, by contacting a
CC biological sample from the patient with a polynucleotide that selectively
CC hybridizes to a sequence at least 80% identical to any of one of 80
CC nucleic acid sequences given in the specification. The method is useful
CC in diagnosing ovarian cancer and in identifying and using agents and/or
CC targets that inhibit ovarian cancer. The nucleic acid molecule
CC polypeptide and the antibody may also be used in detecting ovarian

CC cancers, monitoring and early detection of relapse following treatment,
CC monitoring response to therapy, selecting patients for post-operative
CC chemotherapy or radiation therapy, in selection mode of therapy,
CC determining tumour prognosis, early detection of pre-cancerous lesions,
CC and as vaccines. This sequence corresponds to one of the proteins used
CC for the detection method of the invention.

XX Sequence 253 AA;

Query Match 100.0%; Score 41; DB 7; Length 253;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
Db 1 MARSLLLPL 9

RESULT 13
AAB21326
ID AAB21326 standard; protein; 257 AA.

XX AC AAB21326;
XX DT 02-FEB-2001 (first entry)
XX DE Human HSCCE.
XX KW Human; KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6; HSCCE;
XX human stratum corneum chymotryptic enzyme; kallikrein-like protein;
XX serine protease; cytostatic; cancer; prostrate cancer.

XX Homo sapiens.

XX WO200053776-A2.
XX PD 14-SEP-2000.
XX PF 09-MAR-2000; 2000WO-CA000258.
XX PR 11-MAR-1999; 99US-0124260P.
XX PR 01-APR-1999; 99US-0127388P.
XX PR 21-JUL-1999; 99US-0144919P.
XX PA (MOUN) MOUNT SINAI HOSPITAL.
XX Yousef GM, Diamandis EP;
XX WI; 2000-587440/55.

XX New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L
PT protein mediated disorders, especially cancer.

XX Example 4; Fig 17; 184pp; English.

XX The present sequence is human stratum corneum chymotryptic enzyme
CC (HSCCE), a member of the kallikrein multi-gene family. Kallikreins and
CC kallikrein-like proteins are a subgroup of the serine protease enzyme
CC family. They catalyse the selective cleavage of specific polypeptide
CC precursors to release peptides with potent biological activity. Nucleic
CC acids encoding kallikrein-like proteins KLK-L1, KLK-L2, KLK-L3, KLK-L4,
CC KLK-L5 and KLK-L6 have been isolated. The proteins are useful in the
CC treatment, monitoring and diagnosis of cancers, especially prostate
CC cancer. They can also be used to identify a substance that can associate
CC with or mediate the biological activity of the proteins. Antibodies can
CC be used to treat conditions mediated by the kallikrein-like proteins

XX Sequence 257 AA;

Query Match 100.0%; Score 41; DB 3; Length 257;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
DB 1 MARSLLLPL 9

RESULT 14

AAE08320
ID AAE08320 standard; peptide; 9 AA.
XX
AC AAE08320;
DT 01-NOV-2001 (first entry)
XX
XX Human stratum corneum chymotrypsin enzyme peptide #85 (residues 2-10).
XX
XX Stratum corneum chymotrypsin enzyme; SCCE; cytostatic; vaccine; tumour;
KW cancer; ovarian, breast; lung; colon; prostate; carcinoma; human;
KW antisense therapy; malignant hyperplasia.
XX
OS Homo sapiens.
XX
XX WO200159158-A1.
XX
XX 16-AUG-2001.
XX
XX 07-FEB-2001; 2001WO-US003977.
PF
XX 11-FEB-2000; 2000US-00502600.
PR
XX (UYAR-) UNIV ARKANSAS.
PA
XX
XX O'Brien TJ;
PI
XX WPI; 2001-514676/56.
DR
XX
XX Diagnosing cancer comprises detecting stratum corneum chymotrypsin
PT enzyme.
XX
XX Disclosure; Page 121; 127pp; English.
PS

XX The invention relates to diagnosing cancer especially ovarian cancer, by
CC screening for stratum corneum chymotrypsin enzyme (SCCE). Proteases are
CC considered to be an integral part of tumour growth and metastasis, and
CC therefore, markers indicative of their presence or absence are useful for
CC the diagnosis of cancer. The method is useful for diagnosing cancer,
CC carcinoma and malignant hyperplasia. The SCCE oligonucleotide may be used
CC to treat a cancer selected from ovarian, breast, lung, colon, prostate
CC and other cancers in which SCCE is overexpressed. The present sequence is
CC human SCCE peptide
XX
SQ Sequence 9 AA;
Query Match 87.8%; Score 36; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 ARSLLLPL 9
DB 1 ARSLLLPL 8

RESULT 15

ABB98135
ID ABB98135 standard; protein; 818 AA.
XX
XX ABB98135;
AC
XX
DT 17-OCT-2002 (first entry)
XX
XX Human PMMM Incyte ID 2751509CD1.
DE
XX Human; PMMM; protein modification and maintenance molecule;
KW anticonvulsant; neuroprotective; nootropic; cytosolic; antipsoriatic;
KW

KW antiasthmatic; dermatological; antidiabetic; antiparkinsonian;
KW antianaemic; antiinflammatory; antiulcer; antianginal; cardiac;
KW hepatotropic; osteopathic; antiemetic; antipyretic; virucide;
KW antibacterial; fungicide; gastrointestinal; antidiarrheic; laxative;
KW haemostatic; anti-HIV; thrombolytic; anticoagulant; gynaecological;
KW cardiovascular; antiarteriosclerotic; hypotensive; vasotropic;
KW antitumour; antirheumatic; immunosuppressive; antiallergic; antithyroid;
KW nephrotoxic; antitoxic; thyromimetic; antiarthritic; uropathic;
KW ophthalmological; antiparasitic; tranquiliser; vulnerary; keratolytic;
KW auditory; antiseborrheic; antidepressant; neuroleptic; antiferility;
KW anthelmintic; protozoacide; Crohn's disease; hypertension; autoimmune;
KW inflammatory; anaemia; cell proliferative; developmental; epithelial;
KW scabies; neurological; Alzheimer's disease; reproductive;
KW ectopic pregnancy; gene therapy; vaccine; disorder; prostasin.
XX
XX Homo sapiens.
OS
XX WO200246383-A2.
PN
XX 13-JUN-2002.
PD
XX 05-DEC-2001; 2001WO-US046964.
PF
XX 08-DEC-2000; 2000US-0254399P.
PR
XX 21-DEC-2000; 2000US-0257803P.
PR
XX 05-JAN-2001; 2001US-0260110P.
PR
XX 19-JAN-2001; 2001US-0262851P.
PR
XX 25-JAN-2001; 2001US-0264623P.
PR
XX (INCY-) INCYTE GENOMICS INC.
PA
XX Yue H, Azimzai Y, Kallick DA, Baughn MR, Griffin JA, Swarnakar A;
PI Lal PG, Wallia NK, Hafalia AJA, Gandhi AR, Au-Young J, Elliott VS;
PI Ramkumar J, Thangavelu K, Lu Y, Warren BA, Lu DAM, Lee EA;
PI Tribouley CM, Arvizu C, Delegeane AM, Yao MG, Khan FA;
PI Sanjanwala MM;
XX
XX WPI; 2002-519664/55.
DR
XX N-PSDB; ABQ75956.
DR
XX New isolated Protein Modification and Maintenance polypeptides, useful
PT for diagnosis, and treatment of e.g. gastrointestinal disorders.
PT
XX Claim 1 (a); Page 174-176; 200pp; English.
PS
XX The invention relates to an isolated Protein Modification and Maintenance
CC (PMM) polypeptide. Polypeptides of the invention may be used in the
CC diagnosis, treatment and prevention of disorders associated with
CC decreased expression or activity of PMM. These include gastrointestinal
CC disorders (e.g. Crohn's disease), cardiovascular disorders (e.g.
CC hypertension), autoimmune/inflammatory disorders (e.g. anaemia), cell
CC proliferative disorders, developmental disorders, epithelial disorders
CC (e.g. scabies), neurological disorders (e.g. Alzheimer's disease)
CC reproductive disorders (e.g. ectopic pregnancy), and in gene therapy or a
CC vaccine for such diseases. They may also be used in the assessment of the
CC effects of exogenous compound on the expression of nucleic acid and amino
CC acid sequences of protein modification and maintenance molecules. The
CC current sequence represents a human PMM of the invention, which has been
CC found to have homology with rat prostasin
XX
SQ Sequence 818 AA;

Query Match 87.8%; Score 36; DB 5; Length 818;
Best Local Similarity 88.9%; Pred. No. 1.1e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
DB 1 MARHLLPL 9

Search completed: March 1, 2004, 17:28:56
Job time : 46.5556 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:21:46 ; Search time 11.889 Seconds
(without alignment)

39.081 Million cell updates/sec

Title: US-09-905-083-86

Perfect score: 41

Sequence: 1 MARSLLLPL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/2/iaa/5A COMB.pbp: *
2: /cgn2_6/ptodata/2/iaa/5B COMB.pbp: *
3: /cgn2_6/ptodata/2/iaa/6A COMB.pbp: *
4: /cgn2_6/ptodata/2/iaa/6B COMB.pbp: *
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pbp: *
6: /cgn2_6/ptodata/2/iaa/backfiles1.pbp: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	41	100.0	9	3	US-09-502-600-86
2	41	100.0	9	4	US-09-918-243-86
3	41	100.0	253	2	US-08-557-146-2
4	41	100.0	253	2	US-08-824-874-3
5	41	100.0	253	2	US-09-154-344-2
6	41	100.0	253	3	US-08-930-188-2
7	41	100.0	253	3	US-09-210-084-3
8	41	100.0	253	4	US-09-764-762-3
9	41	100.0	253	5	PCT-US96-04294-2
10	36	87.8	9	3	US-09-502-600-116
11	36	87.8	9	4	US-09-918-243-116
12	35	85.4	518	3	US-08-999-723-2
13	35	85.4	518	3	US-09-434-427-2
14	35	85.4	518	4	US-09-548-372D-2
15	35	85.4	518	4	US-09-548-367D-2
16	35	85.4	518	4	US-09-551-853D-2
17	35	85.4	518	4	US-09-215-450-19
18	32	78.0	43	4	US-09-149-476-355
19	32	78.0	204	4	US-09-134-000C-3659
20	32	78.0	238	4	US-09-149-476-485
21	32	78.0	303	3	US-08-985-950-2
22	32	78.0	303	4	US-09-546-049-2
23	32	78.0	493	4	US-09-252-991A-30722
24	32	78.0	654	4	US-09-907-794A-177
25	32	78.0	654	4	US-09-905-125A-177
26	32	78.0	654	4	US-09-902-775A-177
27	31	75.6	23	1	US-08-353-751-1

28 75.6 190 4 US-09-543-681A-7684 Sequence 7684, Ap
29 31 75.6 319 4 US-09-489-039A-7873 Sequence 7873, Ap
30 31 75.6 2005 3 US-08-836-325-7 Sequence 7, Appli
31 76 4 US-09-252-991A-21761 Sequence 21761, A
32 76 4 US-09-732-210-97 Sequence 97, Appl
33 76 4 US-09-732-210-301 Sequence 301, Appl
34 76 4 US-09-198-119C-93 Sequence 93, Appl
35 76 4 US-09-198-119C-15 Sequence 15, Appl
36 76 4 US-09-301-666A-2 Sequence 2, Appli
37 76 4 US-09-301-217-2 Sequence 2, Appli
38 76 4 US-09-252-991A-25019 Sequence 25019, A
39 76 4 US-09-252-991A-17109 Sequence 17109, A
40 76 4 US-09-489-039A-10804 Sequence 10804, A
41 76 4 US-09-489-039A-9156 Sequence 9156, Ap
42 76 4 US-09-587-811A-2 Sequence 2, Appli
43 76 4 US-09-134-000C-6359 Sequence 6359, Ap
44 76 4 US-08-353-751-2 Sequence 2, Appli
45 76 4 US-09-107-532A-7115 Sequence 7115, Ap

ALIGNMENTS

RESULT 1
US-09-502-600-86
; Sequence 86, Application US/09502600A
; Patent No. 6294344
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP-C
; CURRENT FILING DATE: 2000-02-11
; CURRENT APPLICATION NUMBER: US/09/502,600A
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1996
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 86
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 1-9 of the SCC3 protein
US-09-502-600-86

Query Match 100.0%; Score 41; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9
Db 1 MARSLLLPL 9

RESULT 2
US-09-918-243-86
; Sequence 86, Application US/09918243
; Patent No. 6627403
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 86
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens

```

;
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 1-9 of the SCCE protein
US-09-918-243-86

Query Match          100.0%; Score 41; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e-05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9
Db 1 MARSLLLPL 9

RESULT 3
US-08-557-146-2
; Sequence 2, Application US/08557146
; Patent No. 5834290
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hansson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,146
; FILING DATE: 14-DEC-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-181
; TELEPHONE: (212) 819-8783
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-557-146-2

Query Match          100.0%; Score 41; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9
Db 1 MARSLLLPL 9

RESULT 4
US-08-824-874-3
; Sequence 3, Application US/08824874
; Patent No. 5962300
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NOVEL KALLIKREIN
; NUMBER OF SEQUENCES: 5
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/824,874
; FILING DATE: Filed Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PP-0252 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 532504
US-08-824-874-3

Query Match          100.0%; Score 41; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9
Db 1 MARSLLLPL 9

RESULT 5
US-09-154-344-2
; Sequence 2, Application US/09154344
; Patent No. 5981256
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hansson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case, Patent Department
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/154,344
; FILING DATE: 16-SEP-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/824,874
; FILING DATE: Filed Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PP-0252 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 532504
US-08-824-874-3
```

APPLICATION NUMBER: US 08/557,146
FILING DATE: 14-DEC-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Starnet, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-181
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-154-344-2

Query Match 100.0%; Score 41; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.32; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9
Db 1 MARSLLLPL 9

RESULT 6
US-08-930-188-2
Sequence 2, Application US/08930188
Patent No. 6093397
GENERAL INFORMATION:
APPLICANT: Dixon, Eric P.
APPLICANT: Johnstone, Edward M.
TITLE OF INVENTION: AMYLOID PRECURSOR PROTEIN PROTEASE AND
RELATED NUCLEIC ACIDS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: United States of America
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/930,188
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/416,257
FILING DATE: 04-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Blalock, Donna K.
REGISTRATION NUMBER: 38,082
REFERENCE/DOCKET NUMBER: X9239
TELEPHONE: 317-277-1090
TELEFAX: 317-276-3861
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-930-188-2

Query Match 100.0%; Score 41; DB 3; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.32; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9
Db 1 MARSLLLPL 9

RESULT 7
US-09-210-084-3
Sequence 3, Application US/09210084
Patent No. 6197511
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
TITLE OF INVENTION: NOVEL KALLIKREIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/210,084
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/824,874
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0252 US
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 532504
US-09-210-084-3

Query Match 100.0%; Score 41; DB 3; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.32; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9
Db 1 MARSLLLPL 9

RESULT 8
US-09-764-762-3
Sequence 3, Application US/09764762
Patent No. 6472195
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
TITLE OF INVENTION: NOVEL KALLIKREIN
NUMBER OF SEQUENCES: 5

;;
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Incyte Pharmaceuticals, Inc.
;; STREET: 3174 Porter Drive
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FASTSEQ for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/764,762
;; FILING DATE: 16-Jan-2001
;; CLASSIFICATION: <Unknown>
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 09/210,084
;; FILING DATE: <Unknown>
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Billings, Lucy J.
;; REGISTRATION NUMBER: 36,749
;; REFERENCE/DOCKET NUMBER: PF-0252 US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-855-0555
;; TELEFAX: 415-845-4166
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 253 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; IMMEDIATE SOURCE:
;; LIBRARY: GenBank
;; CLONE: 532504
;; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-764-762-3

Query Match 100.0%; Score 41; DB 4; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.32; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

QY 1 MARSLLLPL 9
Db 1 MARSLLLPL 9
|||||

RESULT 9
PCT-US96-04294-2
; Sequence 2, Application PC/TUS9604294
; GENERAL INFORMATION:
; APPLICANT: Dixon, Eric P.
; APPLICANT: Johnstone, Edward M.
; APPLICANT: Little, Sheila P.
; TITLE OF INVENTION: AMYLOID PRECURSOR PROTEIN PROTEASE AND
; TITLE OF INVENTION: RELATED NUCLEIC ACIDS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: United States of America
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/04294
; FILING DATE:
; CLASSIFICATION:

Query Match 87.8%; Score 36; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0;

QY 2 ARSLLLPL 9
Db 1 ARSLLLPL 8
|||||

RESULT 11
US-09-918-243-116
; Sequence 116, Application US/09918243
; Patent No. 6627403
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13

Query Match 100.0%; Score 41; DB 5; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.32; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

QY 1 MARSLLLPL 9
Db 1 MARSLLLPL 9
|||||

RESULT 10
US-09-502-600-116
; Sequence 116, Application US/09502600A
; Patent No. 6294344
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP-C
; CURRENT FILING DATE: 2000-02-11
; CURRENT APPLICATION NUMBER: US/09/502,600A
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 116
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 2-10 of the SCCE protein
US-09-502-600-116

Query Match 87.8%; Score 36; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0;

QY 2 ARSLLLPL 9
Db 1 ARSLLLPL 8
|||||

RESULT 11
US-09-918-243-116
; Sequence 116, Application US/09918243
; Patent No. 6627403
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13

; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 116
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 2-10 of the SCCE protein
US-09-918-243-116

Query Match 87.8%; Score 36; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05; Indels 0;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;

QY 2 ARSLLLPL 9
Db 1 ARSLLLPL 8

RESULT 12
US-08-999-723-2
; Sequence 2, Application US/08999723A
; Patent No. 6025180
; GENERAL INFORMATION:
; APPLICANT: Powell, David J.
; APPLICANT: Southan, Christopher
; APPLICANT: Chapman, Conrad G.
; APPLICANT: Evans, Joanne R.
; TITLE OF INVENTION: ASP1
; FILE REFERENCE: GH70262
; CURRENT APPLICATION NUMBER: US/08/999,723A
; CURRENT FILING DATE: 1997-10-06
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-999-723-2

Query Match 85.4%; Score 35; DB 3; Length 518;
Best Local Similarity 77.8%; Pred. No. 14; Indels 0;
Matches 7; Conservative 2; Mismatches 0; Gaps 0;

QY 1 MARSLLLPL 9
Db 4 LARALLPL 12

RESULT 13
US-09-434-427-2
; Sequence 2, Application US/09434427
; Patent No. 6162630
; GENERAL INFORMATION:
; APPLICANT: POWELL, DAVID J.
; APPLICANT: SOUTHAN, CHRISTOPHER
; APPLICANT: CHAPMAN, CONRAD G.
; APPLICANT: EVANS, JOANNE R.
; TITLE OF INVENTION: ASP1
; FILE REFERENCE: GH-70262-D1
; CURRENT APPLICATION NUMBER: US/09/434,427
; CURRENT FILING DATE: 1999-11-04
; EARLIER APPLICATION NUMBER: US 08/999,723
; EARLIER FILING DATE: 1997-10-06
; EARLIER APPLICATION NUMBER: UK 9626022.9
; EARLIER FILING DATE: 1996-12-14
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 518
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-434-427-2

Query Match 85.4%; Score 35; DB 3; Length 518;
Best Local Similarity 77.8%; Pred. No. 14; Indels 0;
Matches 7; Conservative 2; Mismatches 0; Gaps 0;

QY 1 MARSLLLPL 9
Db 4 LARALLPL 12

RESULT 14
US-09-548-372D-2
; Sequence 2, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; FILE REFERENCE: 29915/62801
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-372D-2

Query Match 85.4%; Score 35; DB 4; Length 518;
Best Local Similarity 77.8%; Pred. No. 14; Indels 0;
Matches 7; Conservative 2; Mismatches 0; Gaps 0;

QY 1 MARSLLLPL 9
Db 4 LARALLPL 12

RESULT 15
US-09-548-367D-2
; Sequence 2, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; FILE REFERENCE: 29915/6280H
; CURRENT APPLICATION NUMBER: US/09/548,367D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-367D-2

Query Match 85.4%; Score 35; DB 4; Length 518;
Best Local Similarity 77.8%; Fred. No. 14;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
:|:|:|
DB 4 LARALLPL 12

Search completed: March 1, 2004, 17:39:26
Job time : 12.8889 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:35:01 ; Search time 24.1111 Seconds
(without alignments)
78.818 Million cell updates/sec

Title: US-09-905-083-86

Perfect score: 41

Sequence: 1 MARSULLPL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 21153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

```

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09D_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	9	9	US-09-918-243-86
2	41	100.0	9	9	US-09-905-083-86
3	41	100.0	253	9	US-09-888-615-98
4	41	100.0	253	9	US-09-764-762-3
5	41	100.0	253	14	US-10-264-283-90
6	41	100.0	253	15	US-10-295-027-498
7	41	100.0	253	15	US-10-173-999-48
8	36	87.8	9	9	US-09-918-243-116
9	36	87.8	9	9	US-09-905-083-116
10	36	87.8	818	9	US-09-888-615-111
11	35	85.4	518	9	US-09-794-927-2
12	35	85.4	518	9	US-09-795-847-2
13	35	85.4	518	9	US-09-794-743-2
14	35	85.4	518	9	US-09-794-748-2
15	35	85.4	518	9	US-09-794-925-2

```

16 35 85.4 518 9 US-09-215-450-19 Sequence 19, Appli
17 35 85.4 518 9 US-09-681-442-2 Sequence 2, Appli
18 35 85.4 518 9 US-09-978-295A-196 Sequence 196, App
19 35 85.4 518 9 US-09-886-143-2 Sequence 2, Appli
20 35 85.4 518 9 US-09-978-697-196 Sequence 196, App
21 35 85.4 518 9 US-09-978-192A-196 Sequence 196, App
22 35 85.4 518 9 US-09-999-832A-196 Sequence 196, App
23 35 85.4 518 10 US-09-978-189-196 Sequence 196, App
24 35 85.4 518 10 US-09-978-608A-196 Sequence 196, App
25 35 85.4 518 10 US-09-978-585A-196 Sequence 196, App
26 35 85.4 518 10 US-09-978-191A-196 Sequence 196, App
27 35 85.4 518 10 US-09-978-403A-196 Sequence 196, App
28 35 85.4 518 10 US-09-978-564A-196 Sequence 196, App
29 35 85.4 518 10 US-09-999-833A-196 Sequence 196, App
30 35 85.4 518 10 US-09-981-915A-196 Sequence 196, App
31 35 85.4 518 10 US-09-978-824-196 Sequence 196, App
32 35 85.4 518 10 US-09-918-585A-196 Sequence 196, App
33 35 85.4 518 10 US-09-978-423A-196 Sequence 196, App
34 35 85.4 518 10 US-09-978-193A-196 Sequence 196, App
35 35 85.4 518 10 US-09-869-414-2 Sequence 2, Appli
36 35 85.4 518 10 US-09-999-830A-196 Sequence 196, App
37 35 85.4 518 10 US-09-978-757A-196 Sequence 196, App
38 35 85.4 518 10 US-09-978-187B-196 Sequence 196, App
39 35 85.4 518 10 US-09-548-366-2 Sequence 2, Appli
40 35 85.4 518 10 US-09-978-843A-196 Sequence 196, App
41 35 85.4 518 10 US-09-978-375A-196 Sequence 196, App
42 35 85.4 518 10 US-09-978-298A-196 Sequence 196, App
43 35 85.4 518 10 US-09-978-188A-196 Sequence 196, App
44 35 85.4 518 10 US-09-978-681A-196 Sequence 196, App
45 35 85.4 518 10 US-09-978-194A-196 Sequence 196, App

```

ALIGNMENTS

```

RESULT 1
US-09-918-243-86
; Sequence 86, Application US/09918243
; Patent No. US20020142317A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 86
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 1-9 of the SCCE protein
US-09-918-243-86

```

```

Query Match 100.0%; Score 41; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MARSULLPL 9
| | | | | | | |
DB 1 MARSULLPL 9

```

```

RESULT 2
US-09-905-083-86
; Sequence 86, Application US/09905083
; Patent No. US20020146708A1
; GENERAL INFORMATION:

```


APPLICANT: O'Brien, Timothy J.
TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
FILE OF INVENTION: Ovarian Cancer
FILE REFERENCE: D6223CIP/C/Div
CURRENT APPLICATION NUMBER: US/09/905,083
CURRENT FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: US 09/502,600
PRIOR FILING DATE: 2000-02-11
NUMBER OF SEQ ID NOS: 136
SEQ ID NO 86
LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CHAIN
OTHER INFORMATION: Residues 1-9 of the SCCE protein
US-09-905-083-86

Query Match 100.0%; Score 41; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLPL 9
Db 1 MARSLLPL 9

RESULT 3
US-09-888-615-98
Sequence 98, Application US/09888615
Patent No. US20020064856A1
GENERAL INFORMATION:
APPLICANT: PLOWMAN, GREGORY
APPLICANT: WHYTE, DAVID
APPLICANT: CAENEPEEL, SEAN
APPLICANT: CHARYCZAK, GLEN
APPLICANT: MANNING, GERARD
APPLICANT: SUDARSANAM, SUCHA
TITLE OF INVENTION: NOVEL PROTEASES
FILE REFERENCE: 038602/1214
CURRENT APPLICATION NUMBER: US/09/888,615
CURRENT FILING DATE: 2001-06-26
PRIOR APPLICATION NUMBER: 60/214,047
PRIOR FILING DATE: 2000-06-26
NUMBER OF SEQ ID NOS: 150
SOFTWARE: Patent in Ver. 2.1
SEQ ID NO 98
LENGTH: 253
TYPE: PRT
ORGANISM: Homo sapiens
US-09-888-615-98

Query Match 100.0%; Score 41; DB 9; Length 253;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLPL 9
Db 1 MARSLLPL 9

RESULT 4
US-09-764-762-3
Sequence 3, Application US/09764762
Patent No. US20020068341A1
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
Lal, Preeti
TITLE OF INVENTION: NOVEL KALLIKREIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESS: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive

CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/764,762
FILING DATE: 16-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/210,084
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0252 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 532504
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-764-762-3

Query Match 100.0%; Score 41; DB 9; Length 253;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLPL 9
Db 1 MARSLLPL 9

RESULT 5
US-10-264-283-90
Sequence 90, Application US/10264283
Publication No. US2003014494A1
GENERAL INFORMATION:
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.590
CURRENT APPLICATION NUMBER: US/10/264,283
CURRENT FILING DATE: 2002-10-02
NUMBER OF SEQ ID NOS: 111
SOFTWARE: Corixa Invention Disclosure Database
SEQ ID NO 90
LENGTH: 253
TYPE: PRT
ORGANISM: Homo sapiens
US-10-264-283-90

Query Match 100.0%; Score 41; DB 14; Length 253;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLPL 9
Db 1 MARSLLPL 9

```
RESULT 6
US-10-295-027-498
; Sequence 498, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT FILING DATE: 2002-11-13
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; Remaining Prior Application data removed - See File Wrapper or PALM.
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 498
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-498

Query Match 100.0%; Score 41; DB 15; Length 253;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
Db 1 MARSLLLPL 9

RESULT 7
US-10-173-999-48
; Sequence 48, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; FILE OF INVENTION: Methods of Screening for Modulators of Ovarian
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002420US
; CURRENT FILING DATE: 2002-06-17
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US/10/173,999
; PRIOR APPLICATION NUMBER: US 60/299,234
```

```
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-173-999-48

Query Match 100.0%; Score 41; DB 15; Length 253;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
Db 1 MARSLLLPL 9

RESULT 8
US-09-918-243-116
; Sequence 116, Application US/09918243
; Patent No. US20020142317A1
; GENERAL INFORMATION:
; APPLICANT: Cannon, Timothy J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 116
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 2-10 of the SCCE protein
US-09-918-243-116

Query Match 87.8%; Score 36; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ARSLLLPL 9
Db 1 ARSLLLPL 8

RESULT 9
US-09-905-083-116
; Sequence 116, Application US/09905083
; Patent No. US20020146708A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP/C/Div
; CURRENT APPLICATION NUMBER: US/09/905,083
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 09/502,600
; PRIOR FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 116
; LENGTH: 9
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 2-10 of the SCCE protein
US-09-905-083-116
```

```
Query Match      87.8%; Score 36; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2 ARSLLPL 9
   |||||
Db 1 ARSLLPL 8
```

```
RESULT 10
US-09-888-615-111
; Sequence 111, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHITE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 111
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-111
```

```
Query Match      87.8%; Score 36; DB 9; Length 818;
Best Local Similarity 88.9%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 MARSLLPL 9
   |||||
Db 1 MARHLLPL 9
```

```
RESULT 11
US-09-794-927-2
; Sequence 2, Application US/09794927
; Patent No. US20010016324A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; FILE REFERENCE: 28341/6280FG
; CURRENT APPLICATION NUMBER: US/09/794,927
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
```

```
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-927-2
```

```
Query Match      85.4%; Score 35; DB 9; Length 518;
Best Local Similarity 77.8%; Pred. No. 1e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MARSLLPL 9
   |||||
Db 4 LARALLPL 12
```

```
RESULT 12
US-09-795-847-2
; Sequence 2, Application US/09795847
; Patent No. US20010018208A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; FILE REFERENCE: 28341/6280DE
; CURRENT APPLICATION NUMBER: US/09/795,847
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-795-847-2
```

```
Query Match      85.4%; Score 35; DB 9; Length 518;
Best Local Similarity 77.8%; Pred. No. 1e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MARSLLPL 9
   |||||
Db 4 LARALLPL 12
```

```
RESULT 13
US-09-794-743-2
; Sequence 2, Application US/09794743
; Patent No. US20010021391A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
```

;; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
;; TITLE OF INVENTION: USES
;; TITLE OF INVENTION: THEREFOR
;; FILE REFERENCE: 28341/6280BC
;; CURRENT APPLICATION NUMBER: US/09/794,743
;; PRIOR FILING DATE: 2001-02-27
;; PRIOR APPLICATION NUMBER: 09/416,901
;; PRIOR FILING DATE: 1999-10-13
;; PRIOR APPLICATION NUMBER: 60/155,493
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: 09/404,133
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: PCT/US99/20881
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: 60/101,594
;; PRIOR FILING DATE: 1998-09-24
;; NUMBER OF SEQ ID NOS: 73
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 2
;; LENGTH: 518
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-794-743-2

Query Match 85.4%; Score 35; DB 9; Length 518;
Best Local Similarity 77.8%; Pred. No. 1e+02; 0; Indels 0; Gaps 0;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
:|||||
DB 4 LARALLPL 12

RESULT 14
US-09-794-748-2
;; Sequence 2, Application US/09794748
;; Patent No. US2002037315A1
;; GENERAL INFORMATION:
;; APPLICANT: Gurney, Mark E.
;; APPLICANT: Bienkowski, Michael J.
;; APPLICANT: Heinrichson, Robert L.
;; APPLICANT: Parodi, Luis A.
;; APPLICANT: Yan, Riqiang
;; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
;; TITLE OF INVENTION: USES
;; TITLE OF INVENTION: THEREFOR
;; FILE REFERENCE: 28341/6280JL
;; CURRENT APPLICATION NUMBER: US/09/794,748
;; CURRENT FILING DATE: 2001-02-27
;; PRIOR APPLICATION NUMBER: 09/416,901
;; PRIOR FILING DATE: 1999-10-13
;; PRIOR APPLICATION NUMBER: 60/155,493
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: 09/404,133
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: PCT/US99/20881
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: 60/101,594
;; PRIOR FILING DATE: 1998-09-24
;; NUMBER OF SEQ ID NOS: 73
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 2
;; LENGTH: 518
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-794-748-2

Query Match 85.4%; Score 35; DB 9; Length 518;
Best Local Similarity 77.8%; Pred. No. 1e+02; 0; Indels 0; Gaps 0;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
:|||||

DB 4 LARALLPL 12

RESULT 15
US-09-794-925-2
;; Sequence 2, Application US/09794925
;; Patent No. US20020084819A1
;; GENERAL INFORMATION:
;; APPLICANT: Gurney, Mark E.
;; APPLICANT: Bienkowski, Michael J.
;; APPLICANT: Heinrichson, Robert L.
;; APPLICANT: Parodi, Luis A.
;; APPLICANT: Yan, Riqiang
;; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES
;; TITLE OF INVENTION: THEREFOR
;; FILE REFERENCE: 28341/6280HI
;; CURRENT APPLICATION NUMBER: US/09/794,925
;; CURRENT FILING DATE: 2001-02-27
;; PRIOR APPLICATION NUMBER: 09/416,901
;; PRIOR FILING DATE: 1999-10-13
;; PRIOR APPLICATION NUMBER: 60/155,493
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: 09/404,133
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: PCT/US99/20881
;; PRIOR FILING DATE: 1999-09-23
;; PRIOR APPLICATION NUMBER: 60/101,594
;; PRIOR FILING DATE: 1998-09-24
;; NUMBER OF SEQ ID NOS: 73
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 2
;; LENGTH: 518
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-794-925-2

Query Match 85.4%; Score 35; DB 9; Length 518;
Best Local Similarity 77.8%; Pred. No. 1e+02; 0; Indels 0; Gaps 0;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
:|||||
DB 4 LARALLPL 12

Search completed: March 1, 2004, 18:08:53
Job time : 25.1111 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:21:01 ; Search time 10.2222 Seconds
(without alignments)

84.690 Million cell updates/sec

Title: US-09-905-083-86

Perfect score: 41

Sequence: 1 MARSLLLPL 9

Scoring table: BL0SUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	41	100.0	253	2	A53968	serine proteinase
2	34	82.9	461	2	A45394	suppressor protein
3	33	80.5	397	2	B87343	conserved hypothet
4	33	80.5	571	2	E96550	hypothetical prote
5	33	80.5	582	2	I48673	matrix metalloprot
6	33	80.5	582	2	I84471	hypothetical prote
7	32	78.0	81	2	B97856	DNA-binding protei
8	32	78.0	126	2	C82169	hypothetical prote
9	32	78.0	432	2	A83060	conserved membrane
10	32	78.0	506	2	B87102	ionotropic glutama
11	32	78.0	921	2	T51136	probable ligand-ga
12	32	78.0	923	2	F84732	hypothetical prote
13	32	78.0	1628	2	T38055	sodium channel alp
14	31	75.6	138	2	I48107	hypothetical prote
15	31	75.6	147	2	S28698	cell division inhi
16	31	75.6	169	2	B28016	hypothetical prote
17	31	75.6	169	2	A83627	NADH2 dehydrogenas
18	31	75.6	294	2	B86450	probable transamin
19	31	75.6	346	2	T11364	probable transamin
20	31	75.6	375	2	C71917	PTS system, n-acet
21	31	75.6	375	2	H64597	stromelysin 3 (EC
22	31	75.6	452	2	A83734	inulinase (EC 3.2.
23	31	75.6	491	2	JC6197	cytolysin B transp
24	31	75.6	556	1	S31330	sodium channel pro
25	31	75.6	708	2	T43109	rifamycin polyketi
26	31	75.6	2005	2	B25019	corazonin precursor
27	31	75.6	5069	2	T17464	hypothetical prote
28	30	73.2	72	2	JC2384	
29	30	73.2	74	2	T17834	

30	30	73.2	170	2	S43476	histone-like DNA-binding protein
31	30	73.2	216	2	JEO297	DRE/CRT-binding protein
32	30	73.2	216	2	T51830	transcription factor
33	30	73.2	222	2	D82132	hypothetical protein
34	30	73.2	299	2	T17832	hypothetical protein
35	30	73.2	314	2	D85294	transcription activator
36	30	73.2	314	2	T05799	transcription activator
37	30	73.2	323	2	A91179	probable transcription factor
38	30	73.2	323	2	B86025	probable transcription factor
39	30	73.2	323	2	S47741	probable transcription factor
40	30	73.2	325	2	A88452	protein ZG155.4 (inferred)
41	30	73.2	325	2	T25122	hypothetical protein
42	30	73.2	325	2	A97482	hypothetical protein
43	30	73.2	325	2	A12699	conserved hypothetical protein
44	30	73.2	331	2	G87494	hypothetical protein
45	30	73.2	354	2	A48931	transmembrane glycoprotein

ALIGNMENTS

RESULT 1

A53968
serine proteinase SCCE precursor - human
N;Alternate names: stratum corneum chymotryptic enzyme
C;Species: Homo sapiens (man)
C;Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 22-Jun-1999
C;Accession: A53968
R;Hansson, L.; Stroemqvist, M.; Baekman, A.; Wallbrandt, P.; Carlstein, A.; Egelrud, T.
J. Biol. Chem. 269, 19420-19426, 1994
A;Title: Cloning, expression, and characterization of stratum corneum chymotryptic enzyme
A;Reference number: A53968; MUID:94308225; PMID:8034709
A;Accession: A53968
A;Status: Preliminary
A;Molecule type: mRNA
A;Residues: 1-253 <MAN>
A;Cross-references: GB:I33404; NID:9521214; PIDN:AAC37551.1; PID:95325504
C;Genetics:
A;Gene: GDB:PRSS6; SCCE
A;Cross-references: GDB:377730
A;Map position: 7q35-7q35
C;Superfamily: trypsin; trypsin homology
F;30-245/Domain: trypsin homology <tr>

Query Match 100.0%; Score 41; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9
|||
Db 1 MARSLLLPL 9

RESULT 2

A46394
suppressor protein SSL1 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein L131; protein YLR005W
C;Species: Saccharomyces cerevisiae
C;Date: 18-May-1994 #sequence_revision 19-Jul-1996 #text_change 11-Aug-2003
C;Accession: A46394; S64827
R;Yoon, H.; Miller, S.P.; Pabich, E.K.; Donahue, T.F.
Genes Dev. 6, 2463-2477, 1992
A;Title: SSL1, a suppressor of a HIS4 5'-UTR stem-loop mutation, is essential for transla
A;Reference number: A46394; MUID:94040711; PMID:1340463
A;Accession: A46394
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-461 <YOO>
A;Cross-references: GB:Z17385; NID:92695; PID:92696
R;Vandenbol, M.; Portetelle, D.; Hilger, F.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64742
A;Accession: S64827

A:Molecule type: DNA
A:Residues: 1-461 <VAN>
A:Cross-references: EMBL:Z73177; NID:g1360293; PID:g1360294; MIPS:YLR005W
A>Note: experimental_source strain S288C
C:Genetics:
A:Gene: SGD:SS11
A:Cross-references: SGD:S0003995; MIPS:YLR005W
A:Map position: 12R
C:Superfamily: RNA polymerase II transcription initiation/nucleotide excision repair fac
C:Keywords: transmembrane protein
F:356-372/Domain: transmembrane #status predicted <TM>

Query Match 82.9%; Score 34; DB 2; Length 461;
Best Local Similarity 77.8%; Pred. No. 21;
Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY 1 MARSLLPL 9
DB 216 MARGLLPLV 224

RESULT 3
B87343
Conserved hypothetical protein CC0757 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: B87343
R:Nickman, W.C.; Feldblym, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBov, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolod
n, J.; Zmolava, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4138-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: B87343
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-397 <STO>
A:Cross-references: GB:AE005673; NID:g13421992; PIDN:AAK22742.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC0757

Query Match 80.5%; Score 33; DB 2; Length 397;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLPL 9
DB 285 LARALLPL 293

RESULT 4
E96550
Hypothetical protein Flm15.13 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: E96550
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, T.H.; Dwar, K.
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E96550
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-571 <STO>
A:Cross-references: GB:AE005173; NID:g4836937; PIDN:AAD30639.1; GSPDB:GN00141
C:Genetics:

A:Gene: F11M15.13
A:Map position: 1

Query Match 80.5%; Score 33; DB 2; Length 571;
Best Local Similarity 77.8%; Pred. No. 44;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MARSLLPL 9
DB 332 MLRSLLVFL 340

RESULT 5
I48673
Matrix metalloproteinase (EC 3.4.24.-) membrane type precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 22-Jun-1999
C:Accession: I48673
R:Okada, A.; Bellocq, J.P.; Rouyer, N.; Chenard, M.P.; Rio, M.C.; Chambon, P.; Basset, P.
Proc. Natl. Acad. Sci. U.S.A. 92, 2730-2734, 1995
A:Title: Membrane-type matrix metalloproteinase (MT-MMP) gene is expressed in stromal cel
A:Reference number: I38046; MUID:95224014; PMID:7708715
A:Accession: I48673
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-582 <RES>
A:Cross-references: EMBL:X83536; NID:G804999; PIDN:CAA58520.1; PID:G805000
C:Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metalloproteir
C:Keywords: hydrolase; metalloproteinase; zinc; zymogen
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-97/Domain: activation peptide #status predicted <PRO>
F:51-284/Domain: matrix metalloproteinase homology <MMP>
F:313-508/Domain: hemopexin repeat homology <PXN>
F:533-239,243,249/Binding site: zinc, catalytic (His)
F:239,243,249/Binding site: zinc, catalytic (His) (active) #status predicted
F:240/Active site: Glu #status predicted

Query Match 80.5%; Score 33; DB 2; Length 582;
Best Local Similarity 87.5%; Pred. No. 45;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ARSLLPL 9
DB 8 SRSLLPL 15

RESULT 6
I84471
Matrix metalloproteinase (EC 3.4.24.-) membrane type precursor - rat
N:Alternate names: membrane-type metalloproteinase
C:Species: Rattus norvegicus (Norway rat)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 22-Jun-1999
C:Accession: I84471, I61946
R:Okada, A.; Bellocq, J.P.; Rouyer, N.; Chenard, M.P.; Rio, M.C.; Chambon, P.; Basset, P.
Proc. Natl. Acad. Sci. U.S.A. 92, 2730-2734, 1995
A:Title: Membrane-type matrix metalloproteinase (MT-MMP) gene is expressed in stromal cel
A:Reference number: I38046; MUID:95224014; PMID:7708715
A:Accession: I84471
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-582 <RES>
A:Cross-references: EMBL:X83537; NID:G805012; PIDN:CAA58521.1; PID:G805013
A:Accession: I61946
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-67, 'M', '69-254', 'A', '256-582 <RE2>
A:Cross-references: EMBL:X91785; NID:g1001926; PIDN:CAA62897.1; PID:g1001927
C:Genetics:
A:Gene: mt-mmp
C:Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metalloproteir
C:Keywords: hydrolase; metalloproteinase; zinc; zymogen
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-97/Domain: activation peptide #status predicted <PRO>

F:61-284/Domain: matrix metalloproteinase homology <MMP>

F:713-508/Domain: hemopexin repeat homology <PXN>
 F:933,239,243,249/Binding site: zinc, catalytic (Cys, His, His) (inhibited) #status
 F:239,243,249/Binding site: zinc, catalytic (His) (active) #status predicted
 F:240/Active site: Glu #status predicted

Query Match 80.5%; Score 33; DB 2; Length 582;
 Best Local Similarity 87.5%; Pred. No. 45;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ARSLLLPL 9
 Db 8 SRSLLLPL 15

RESULT 7

Hypothetical protein RC1250 [imported] - Rickettsia conorii (strain Malish 7)
 C:Species: Rickettsia conorii
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001
 C:Accession: B97856

R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Rickettsia conorii
 Science 293, 2033-2038, 2001
 A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
 A:Reference number: A97700; MUID:21442074; PMID:11557893

A:Accession: B97856
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-81 <KUR>
 A:Cross-references: GB:AE006914; PIDN:AAL03788.1; PID:gl5620385; GSPDB:GN00173
 C:Genetics:
 A:Gene: RC1250

Query Match 78.0%; Score 32; DB 2; Length 81;
 Best Local Similarity 75.0%; Pred. No. 9.3;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ARSLLLPL 9
 Db 43 ARSLLLPI 50

RESULT 8

C82169
 DNA-binding protein inhibitor Id-2-related protein VC1696 [imported] - Vibrio cholerae
 C:Species: Vibrio cholerae
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C:Accession: C82169

R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
 Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, R.
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301
 A:Accession: C82169
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-126 <HEI>
 A:Cross-references: GB:AE004247; GB:AE003852; NID:g9656204; PIDN:AAF94846.1; GSPDB:GN001
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VC1696

A:Map position: 1

Query Match 78.0%; Score 32; DB 2; Length 126;
 Best Local Similarity 87.5%; Pred. No. 15;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MARSLLLP 8
 Db 33 MLRSLLLP 40

RESULT 9

AR3060
 Hypothetical protein PA4684 [imported] - Pseudomonas aeruginosa (strain PAO1)
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: AR3060

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lam,
 ; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
 A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: AR3060
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-432 <STO>
 A:Cross-references: GB:AE004882; GB:AE004091; NID:g9950939; PIDN:AAG08071.1; GSPDB:GN001
 A:Experimental source: strain PAO1
 C:Genetics:
 A:Gene: PA4684

Query Match 78.0%; Score 32; DB 2; Length 432;
 Best Local Similarity 87.5%; Pred. No. 54;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ARSLLLPL 9
 Db 245 ARELLPL 252

RESULT 10

B87102
 conserved membrane protein ML1544 [imported] - Mycobacterium leprae
 C:Species: Mycobacterium leprae

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
 C:Accession: B87102
 R:Coile, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
 R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
 eam, M.A.; Rutherford, K.M.
 Nature 409, 1007-1011, 2001
 A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq

A:Title: Massive gene decay in the leprosy bacillus.

A:Reference number: A86909; MUID:21128732; PMID:11234002

A:Accession: B87102
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-506 <STO>
 A:Cross-references: GS:AL450380; NID:gl3093364; PIDN:CAC30495.1; GSPDB:GN00147
 C:Genetics:
 A:Gene: ML1544
 C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv1782

Query Match 78.0%; Score 32; DB 2; Length 506;
 Best Local Similarity 77.8%; Pred. No. 63;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MARSLLLPL 9
 Db 223 MNRSLLLPL 231

RESULT 11

T51136
 ionotropic glutamate receptor glr5 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 28-Jul-2000 #sequence_revision 28-Jul-2000 #text_change 28-Jul-2000

C:Accession: T51136

R:Davenport, R.J.; Kiegle, E.A.; Tester, M.

submitted to the EMBL Data Library, December 1999

A:Description: GLR5, an ionotropic glutamate receptor ortholog from Arabidopsis.

A:Reference number: 225309

A:Accession: T51136

A:Status: preliminary; translated from GE/EMBL/DBJ

A:Molecule type: mRNA
 A:Residues: 1-921 <DAV>
 A:Cross-references: EMBL:AF210701; PIDN:AAF21042.1
 A:Experimental source: cultivar Columbis
 C:Genetics:
 A:Gene: glt5
 A:Map position: 2

Query Match 78.0%; Score 32; DB 2; Length 921;
 Best Local Similarity 77.8%; Pred. No. 1.2e+02;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
 |||||
 DB 705 MARSLVPL 715

RESULT 12

F84732
 probable ligand-gated ion channel subunit [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: F84732
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tailon, L.;
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: R84420; MUID:20083487; PMID:10617197
 A:Accession: F84732
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-923 <STO>
 A:Cross-references: GB:AE002093; NID:G3831456; PIDN:AAC69938.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2G32400
 A:Map position: 2

Query Match 78.0%; Score 32; DB 2; Length 923;
 Best Local Similarity 77.8%; Pred. No. 1.2e+02;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
 |||||
 DB 707 MARSLVPL 715

RESULT 13

T38055
 hypothetical protein SPAC22F3.14c - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T38055; T38177; S62429
 R:lye, G.; Churcher, C.M.; Barrall, B.G.; Rajandream, M.A.; Walsh, S.V.
 submitted to the EMBL Data Library, February 1995
 A:Reference number: Z21765
 A:Accession: T38055
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-26 <LY1>
 A:Cross-references: EMBL:Z69239; PIDN:CAA93223.1; GSPDB:GN00066; SPDB:SPAC1D4.14
 R:lye, G.; Churcher, C.M.; Barrall, B.G.; Rajandream, M.A.; Walsh, S.V.
 submitted to the EMBL Data Library, October 1995
 A:Reference number: Z21776
 A:Accession: T38177
 A:Molecule type: DNA
 A:Residues: 8-1621 <LY2>
 A:Cross-references: EMBL:Z54285; NID:gi1008429; PIDN:CAA91079.2; GSPDB:GN00066; SPDB:SPAC

C:Genetics:
 A:Gene: SPAC1D4.14
 A:Map position: 1L

Query Match 78.0%; Score 32; DB 2; Length 1628;

Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 RSLLLPL 9
 |||||
 DB 583 RSLLLPL 589

RESULT 14

I48107
 sodium channel alpha subunit - long-tailed hamster (fragment)
 C:Species: Cricetulus longicaudatus (long-tailed hamster)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 20-Aug-1999
 C:Accession: I48107
 R:Lalik, P.H.; Krafte, D.S.; Ciccarelli, R.B.
 Am J Physiol 264, 803-809, 1993
 A:Title: Characterization of endogenous Sodium channel gene expressed in chinese hamster
 A:Reference number: I48107
 A:Accession: I48107
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-138 <RES>
 A:Cross-references: GB:M87540; NID:g191067; PIDN:AAA36978.1; PID:g191068
 C:Genetics:
 A:Gene: chol
 C:Superfamily: sodium channel protein
 C:Keywords: duplication

Query Match 75.6%; Score 31; DB 2; Length 138;
 Best Local Similarity 75.0%; Pred. No. 27;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLP 8
 |||||
 DB 1 MARSLVLP 8

RESULT 15

S28698
 hypothetical protein 16 - Agrobacterium tumefaciens plasmid pTi15955

C:Species: Agrobacterium tumefaciens
 C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 08-Oct-1999
 C:Accession: S28698
 R:Barker, R.F.; Idler, K.B.; Thompson, D.V.; Kemp, J.D.
 Plant Mol. Biol. 2, 335-350, 1983
 A:Title: Nucleotide sequence of the T-DNA region from the Agrobacterium tumefaciens octoI

A:Reference number: S28683
 A:Accession: S28698

A:Status: translation not shown
 A:Molecule type: DNA

A:Residues: 1-147 <BAR>
 A:Cross-references: EMBL:X00493; NID:g39062; PIDN:CAA25178.1; PID:g39078

C:Genetics:
 A:Genome: plasmid

Query Match 75.6%; Score 31; DB 2; Length 147;
 Best Local Similarity 87.5%; Pred. No. 28;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MARSLLLP 8
 |||||
 DB 1 MARVLLLP 8

Search completed: March 1, 2004, 17:36:32
 Job time : 11.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:17:25 ; Search time 6 Seconds
(without alignments)
78.105 Million cell updates/sec

Title: US-09-905-083-86
Perfect score: 41
Sequence: 1 MARSLLLPL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	253	1	KLK7 HUMAN
2	35	85.4	518	1	BAE22 HUMAN
3	34	82.9	461	1	SSL1 YEAST
4	33	80.5	464	1	OTSA RHISN
5	33	80.5	582	1	MM14 MOUSE
6	33	80.5	582	1	MM14 RAT
7	32	78.0	580	1	MM14 PIG
8	32	78.0	968	1	SNXD HUMAN
9	32	78.0	1628	1	YATE SCHPO
10	31	75.6	169	1	SULA SALTY
11	31	75.6	346	1	NU2M EQUAS
12	31	75.6	2005	1	CIN2 RAT
13	30	73.2	148	1	CYTC RABIT
14	30	73.2	154	1	COR2 DROME
15	30	73.2	299	1	YHUC ECOLI
16	30	73.2	354	1	CD68 HUMAN
17	30	73.2	357	1	CVHY GIBBA
18	30	73.2	575	1	PT1 ECOLI
19	30	73.2	575	1	PT1 HAEIN
20	30	73.2	575	1	PT1 SALTY
21	30	73.2	1318	1	VIVD EPT7
22	30	73.2	2554	1	7LES DROME
23	29	70.7	101	1	PIPI PIG
24	29	70.7	129	1	IGF2 MUSVI
25	29	70.7	167	1	QCR4 BACSV
26	29	70.7	222	1	IF3B DROME
27	29	70.7	247	1	Y284 PYRHO
28	29	70.7	247	1	Y311 PYRPU
29	29	70.7	260	1	NIRQ PSEAE
30	29	70.7	276	1	PVGI SYNEL
31	29	70.7	328	1	TH12 SCHPO
32	29	70.7	370	1	CLCB PSEPU
33	29	70.7	463	1	HSJU THEMA

34	29	70.7	555	1	INUL KLUMA
35	29	70.7	705	1	MMLD STRCO
36	29	70.7	730	1	CATA HALMA
37	29	70.7	992	1	PHS2 DICDI
38	29	70.7	1247	1	NIDC HUMAN
39	29	70.7	1429	1	EXPA DROME
40	29	70.7	3564	1	CSM1 MOUSE
41	28	68.3	105	1	YA83 HAEIN
42	28	68.3	191	1	RL9 RHIME
43	28	68.3	192	1	RL9 RHILT
44	28	68.3	192	1	Y857 METH
45	28	68.3	221	1	IFB3 ARATH

ALIGNMENTS

RESULT 1	KLK7_HUMAN	STANDARD;	PRT;	253 AA.
ID	KLK7_HUMAN			
AC	P49862;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DE	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Kallikrein 7 precursor (EC 3.4.21.-) (Stratum corneum chymotryptic enzyme) (hSCCE).			
DE	KLK7 OR PRSS6 OR SCCE.			
GN	Homo sapiens (Human).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 23-53.			
RC	TISSUE=Skin;			
RX	MEDLINE=94308225; PubMed=8034709;			
RA	Hansson L., Stromqvist M., Baekman A., Wallbrandt P., Carlstein A., Egelrud T.;			
RT	"Cloning, expression, and characterization of stratum corneum chymotryptic enzyme. A skin-specific human serine proteinase.";			
RL	J. Biol. Chem. 269:19420-19426(1994).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Keratinocytes;			
RA	Yousef G.M., Scorrilas A., Diamandis E.P.;			
RT	"Molecular characterization, mapping and tissue expression of the human stratum corneum chymotryptic enzyme gene.";			
RL	Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20510030; PubMed=11054574;			
RA	Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McQuaig J., Moss P., Paepser B., Wang K.;			
RT	"Sequencing and expression analysis of the serine protease gene cluster located in chromosome 19q13 region.";			
RL	Gene 257:119-130(2000).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	Hansson L., Baekman A., Ny A., Edlund M., Edholm E., Tornell J., Wallbrandt P., Egelrud T.;			
RT	"Epidermal overexpression of stratum corneum chymotryptic enzyme in mice: a model for chronic itchy dermatitis.";			
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	CHARACTERIZATION.			
RX	MEDLINE=95314630; PubMed=7794273;			
RA	Skytt A., Stromqvist M., Egelrud T.;			
RT	"Primary substrate specificity of recombinant human stratum corneum chymotryptic enzyme.";			
RL	Biochem. Biophys. Res. Commun. 211:586-589(1995).			
CC	-!- FUNCTION: May catalyze the degradation of intercellular cohesive structures in the cornified layer of the skin in the continuous shedding of cells from the skin surface. Specific for amino acid residues with aromatic side chains in the P1 position. SCCE			

CC cleaves insulin B chain at 6-Leu-|-Cys-7, 16-Tyr-|-Leu-17, 25-Phe-
 CC -|-Tyr-26, and 26-Tyr-|-Thr-27. Could play a role in the activation
 CC of precursors to inflammatory cytokines.
 CC -|- TISSUE SPECIFICITY: Abundantly expressed in the skin and is
 CC expressed by keratinocytes in the epidermis. Very low levels are
 CC also seen in the brain and kidney.
 CC -|- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L33404; AAC37551.1; --
 DR EMBL: AF166330; AAD49718.1; --
 DR EMBL: AF243527; AAG33360.1; --
 DR EMBL: AF332583; AAK69624.1; --
 DR PIR: A53968; A53968;
 DR HSSP: P00763; IDPO.
 DR MEROPS: S01.300; --
 DR Genem: HGNC:6368; KLK7.
 DR MIM: 604438; --
 DR GO: GO:0008236; F:serine-type peptidase activity; TAS.
 DR GO: GO:0008544; P:epidermal differentiation; TAS.
 DR InterPro: IPR009003; Cys_Ser_cryptin.
 DR InterPro: IPR001254; Peptidase_S1.
 DR InterPro: IPR001314; Peptidase_S1A.
 DR Pfam: PF00089; trypsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SMC0020; TRYPSIN; 1.
 DR PROSITE: PS00240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 DR HydroLase: Serine protease; Zymogen; Glycoprotein; Signal.
 KW SIGNAL 1 22
 FT PROPEP 23 29 ACTIVATION PEPTIDE.
 FT CHAIN 30 253 KALLIKREIN 7.
 FT ACT_SITE 70 70 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 112 112 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 205 205 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 36 137 BY SIMILARITY.
 FT DISULFID 55 71 BY SIMILARITY.
 FT DISULFID ? 239 BY SIMILARITY.
 FT DISULFID 144 211 BY SIMILARITY.
 FT DISULFID 176 190 BY SIMILARITY.
 FT DISULFID 201 226 BY SIMILARITY.
 FT CARBOHYD 246 N-LINKED (GLCNAC...) (POTENTIAL).
 FT SEQUENCE 253 AA; 27525 MW; 2D68B6B15A76A668 CRC64;
 Query Match 100.0%; Score 41; DB 1; Length 253;
 Best Local Similarity 100.0%; Pred. NO. 0.19;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MARSLLPL 9
 Db 1 MARSLLPL 9
 RESULT 2
 BAE2_HUMAN STANDARD; PRT; 518 AA.
 ID BAE2_HUMAN
 AC Q9Y5Z0; Q9UT6;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Beta secretase 2 precursor (EC 3.4.23.45) (Beta-site APP-cleaving
 DE enzyme 2) (Aspartyl protease 1) (Asp 1) (ASP1) (Membrane-associated
 DE aspartic protease 1) (Memapsin-1) (Down region aspartic protease).
 GN BACE2 OR ASP21.
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20057170; PubMed=10591213;
 RA Yan R., Bienkowski M.J., Shuck M.E., Miao H., Tory M.C., Pauley A.M.,
 RA Brashler J.R., Stratman N.C., Mathews W.R., Ruhl A.B., Carter D.B.,
 RA Tomaselli A.G., Parodi L.A., Heinrikson R.L., Gurney M.E.;
 RT "Membrane-anchored aspartyl protease with Alzheimer's disease
 RT beta-secretase activity."; Nature 402:533-537(1999).
 RL Nature 402:533-537(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Bone marrow;
 RA Xin H., Stephens J.C., Duan X., Harrowe G., Kim E., Grieshammer U.,
 RA Giese K.;
 RT "Identification of a novel aspartic-like protease differentially
 RT expressed in human breast cancer cell lines."; Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX Accarino M.P., Fumagalli P., Ottolenghi S., Taramelli R.;
 RA "Cloning of a gene from chromosome 21 Down region encoding a potential
 RA transmembrane aspartyl protease."; Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RX Solans A., Bativill X., de la Luna S.;
 RA "Cloning of a novel mammalian aspartyl protease."; Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20120043; PubMed=10656250;
 RA Hussain I., Powell D.J., Howlett D.R., Tew D.G., Meek T.D.,
 RA Chapman C., Gloger I.S., Murphy K.E., Southan C.D., Ryan D.M.,
 RA Smith T.S., Simmons D.L., Walsh F.S., Dingwall C., Christie G.;
 RT "Identification of a novel aspartic proteinase (Asp 2) as
 RT beta-secretase."; Mol. Cell. Neurosci. 14:419-427(1999).
 RL Mol. Cell. Neurosci. 14:419-427(1999).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20144060; PubMed=10677483;
 RA Lin X., Koelsch G., Wu S., Downs D., Dashti A., Tang J.;
 RT "Human aspartic protease memapsin 2 cleaves the beta-secretase site of
 RT beta-amyloid precursor protein."; Proc. Natl. Acad. Sci. U.S.A. 97:1456-1460(2000).
 RL Proc. Natl. Acad. Sci. U.S.A. 97:1456-1460(2000).
 RN [7]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20289799; PubMed=10830953;
 RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
 RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y.,
 RA Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,
 RA Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
 RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
 RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
 RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
 RA Minoshima S., Shimizu N., Nordieck G., Hornischer K., Brandt P.,
 RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
 RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
 RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
 RA Lehrach H., Reinhardt R., Yaspo M.-L.;
 RT "The DNA sequence of human chromosome 21."; Nature 405:311-319(2000).
 RL Nature 405:311-319(2000).
 RN [8]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Skin;
 RA MEDLINE=22389257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedon T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Iquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahay J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903 (2002).
RP CHARACTERIZATION.
RX MEDLINE=2208158; PubMed=12093293;
RA Turner R.T. III, Loy J.A., Nguyen C., Devasamudram T., Ghosh A.K.,
RA Koelsch G., Tang J.,
RT "Specificity of memapsin 1 and its implications on the design of
RA menapsin 2 (beta-secretase) inhibitor selectivity";
RL Biochemistry 41:8742-8746 (2002).
CC -!- CATALYTIC ACTIVITY: Broad endopeptidase specificity. Cleaves Glu-
CC Val-Asn-Leu-|-Asp-Ala-Glu-Phe in the Swedish variant of
CC Alzheimer's amyloid precursor protein.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Belongs to peptidase family A1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
DR EMBL; AF200342; AAF17078.1; -
DR EMBL; AF200342; AAD45240.1; -
DR EMBL; AF050171; AAD45963.1; -
DR EMBL; AF178532; AAF29494.1; -
DR EMBL; AF204944; AAF26369.1; -
DR EMBL; AF200192; AAF13714.1; -
DR EMBL; AL163284; CAB90458.1; -
DR EMBL; AL163285; CAB90554.1; -
DR EMBL; BC044453; AAH44453.1; -
DR HSP; P00797; 2REN.
DR MEROPS; A01.041; -
DR Genew; HGNC:934; BACE2.
DR MIM; 605668; -
DR GO; GO:0005624; C:membrane fraction; TAS.
DR GO; GO:0004120; P:aspartic-type endopeptidase activity; TAS.
DR GO; GO:0006464; P:protein modification; TAS.
DR GO; GO:0009306; P:protein secretion; TAS.
DR InterPro; IPR001969; Asparticase_AS.
DR InterPro; IPR009007; Pept_Acid.
DR InterPro; IPR001461; Peptidase_A1.
DR Pfam; PF00026; asp; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
KW Hydrolase; Aspartyl protease; Glycoprotein; Zymogen; Transmembrane;
Signal.
FT SIGNAL 1 20 POTENTIAL.
FT PROPEP 21 ? POTENTIAL.
FT CHAIN ? 518 BETA SECRETASE 2.
FT DOMAIN 21 473 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 474 494 POTENTIAL.
FT TRANSMEM 495 518 CYTOPLASMIC (POTENTIAL).
FT ACT_SITE 110 110 BY SIMILARITY.
FT ACT_SITE 303 303 BY SIMILARITY.
FT CARBOHYD 170 170 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 365 365 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 36 A -> T (IN REF. 6).
SQ SEQUENCE 518 AA; 56180 MW; 2E903150823760D3 CRC64;

Query Match 85.4%; Score 35; DB 1; Length 518;
Best Local Similarity 77.8%; Pred. No. 7.6;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MARSLLLP 9
DB 4 LARALLLP 12
RESULT 3
SSLI_YEAST STANDARD; PRT; 461 AA.
ID SSLI_YEAST
AC Q04673;
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Suppressor of stem-loop protein 1.
GN SSLI OR YLR005W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID:4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94040711; PubMed=1340463;
RY Yoon H., Miller S.P., Pabich E.K., Donahue T.F.;
RT "SSLI, a suppressor of a His4 5'-UTR stem-loop mutation, is essential
RL for translation initiation and affects UV resistance in yeast";
RL Genes Dev. 6:2463-2477 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=8288c / AB972;
MEDLINE=97313367; PubMed=9169871;
RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansonge W.,
RA Benes V., Bruckner M., Delius H., Dubois E., Duesterhoeft A.,
RA Entian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,
RA Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Koetter P.,
RA Louis E.J., Messenguy F., Mewes H.-W., Miosga T., Moestl D.,
RA Mueller-Auer S., Nentwich U., Obermaier B., Piravandi E., Pohl T.M.,
RA Portelle D., Purnelle B., Reckmann S., Rieger M., Rinke M., Rose M.,
RA Schaffe M., Scherens B., Scholler P., Schwager C., Schwarz S.,
RA Underwood A.P., Urrestazu L.A., Vandenbol M., Verhaeselt P.,
RA Vierendeels F., Voet M., Voickaert G., Voss H., Wambutt R., Wedler E.,
RA Wedler H., Zimmermann F.K., Zollner A., Hani J., Hoheisel J.D.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII";
RL Nature 387:87-90 (1997).
RN [3]
RP SEQUENCE OF 400-461 FROM N.A.
RX MEDLINE=94239498; PubMed=8183345;
RA Maeda T., Wurgler-Murphy S.M., Saito H.;
RT "A two-component system that regulates an osmosensing MAP kinase
RL cascade in yeast";
RL Nature 369:242-245 (1994).
CC -!- FUNCTION: ESSENTIAL FOR TRANSLATION INITIATION AND AFFECTS
CC UV-RESISTANCE IN YEAST.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: Contains 1 C2H2-type zinc finger.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
DR EMBL; Z17385; CAA78992.1; -
DR EMBL; Z17377; CAA97527.1; -
DR EMBL; L28523; AAA35101.1; -
DR F1R; A46394; A46394.
DR GeneOnline; 142067; -
DR TRANSFAC; T02191; -

```

DR SGD; S0003995; S5L1.
DR InterPro; IPR004595; S5L1.
DR InterPro; IPR007138; S5L1 like.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF04036; S5L1; 1.
DR Pfam; PF00096; Zf-C2H2; 1.
DR SMART; SM00355; Znf_C2H2; 1.
DR TIGRfam; TIGR00622; S5L1; 1.
DR PROSITE; PS00028; ZINC FINGER C2H2 1; 1.
DR PROSITE; PS00157; ZINC FINGER C2H2 2; FALSE NEG.
DR Zinc-finger; Nuclear protein; Metal-binding.
KW ZNF_FING 427 449
FT ZNF_FING 427 449
SQ SEQUENCE 461 AA; 5290 MW; 57ADCB630B790B4F CRC64;

Query Match 82.9%; Score 34; DB 1; Length 461;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 0; Gaps 0;

Qy 1 MARSLILPL 9
Db 216 MARGLLLPV 224

RESULT 4
ID OTSA_RHSN STANDARD; PRT; 464 AA.
AC P55612;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable alpha, alpha-trehalose-phosphate synthase [UDP-forming]
DE (EC 2.4.1.15) (Trehalose-6-phosphate synthase) (UDP-glucose-
DE glucosephosphate glucosyltransferase).
GN OTSA OR Y4PC. (strain NGR234).
OS Rhizobium sp. (strain NGR234).
OG Plasmid sym pNGR234.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Fellay R., Baistroch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -!- CATALYTIC ACTIVITY: UDP-glucose + D-glucose 6-phosphate = UDP +
CC alpha, alpha-trehalose 6-phosphate.
CC -!- SIMILARITY: Belongs to the glycosyltransferase family 20.
CC -!- SIMILARITY: STRONG, TO A NON-FUNCTIONAL COPY IN NGR234, FQ1-FQ2,
CC TRUNCATED BY A ISRM3-LIKE INSERTION ELEMENT.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AF000090; AAB91813.1;
CC InterPro; IPR001830; Glyco trans 20.
CC Pfam; PF00582; Glyco trans 20; 1.
CC Transferase; Glycosyltransferase; Plasmid.
KW TRANSFERASE
SQ SEQUENCE 464 AA; 51627 MW; 96DC610DD739FA73 CRC64;

Query Match 80.5%; Score 33; DB 1; Length 464;
Best Local Similarity 77.8%; Pred. No. 18;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MARSLILPL 9
Db 418 MARSLIMPL 426

```

RESULT 5
 NM14 MOUSE STANDARD; PRT; 582 AA.
 ID P53690; O08645; O35369;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Matrix metalloproteinase-14 precursor (EC 3.4.24.80) (MMP-14)
 DE (Membrane-type 1 matrix metalloproteinase 1) (MT1-MMP 1) (MTMMP1)
 DE (Membrane-type 1 matrix metalloproteinase) (MT1-MMP) (MTMMP)
 DE X1 (MT-MMP).
 GN MMP14 OR MTMMP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95224014; PubMed=7708715;
 RA Okada A., Bellocq J.P., Rouyer N., Chenard M.P., Rio M.C.,
 RA Chambon P., Basset P.;
 RT "Membrane-type matrix metalloproteinase (MT-MMP) gene is expressed in
 RT stromal cells of human colon, breast, and head and neck carcinomas.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:2730-2734(1995).
 RN [2]
 RP REVISIONS.
 RA Osaka A.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SV;
 RX MEDLINE=97467335; PubMed=9325265;
 RA Apte S.S., Fukui N., Beier D.R., Olsen B.R.;
 RT "The matrix metalloproteinase-14 (MMP-14) gene is structurally
 RT distinct from other MMP genes and is co-expressed with the TIMP-2 gene
 RT during mouse embryogenesis";
 RL J. Biol. Chem. 272:25511-25517(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CD-1; TISSUE=Kidney;
 RX MEDLINE=98311877; PubMed=9648071;
 RA Ota K., Stetler-Stevenson W.G., Yang Q., Kumar A., Wada J.,
 RA Kashiwara N., Wallner E.I., Kanwar Y.S.;
 RT "Cloning of murine membrane-type-1 matrix metalloproteinase (MT-1-MMP)
 RT and its metaphoric developmental regulation with respect to MMP-2
 RT and its inhibitor.";
 RL Kidney Int. 54:131-142(1998).
 RN [5]
 RP FUNCTION
 RX MEDLINE=99449306; PubMed=10520996;
 RA Holmbeck K., Bianco P., Caterina J., Yamada S., Kromer M.,
 RA Kuznetsov S.A., Mankani M., Robey P.G., Poole A.R., Pidoux I.,
 RA Ward J.M., Birkedal-Hansen H.;
 RT "MT1-MMP-deficient mice develop dwarfism, osteopenia, arthritis, and
 RT connective tissue disease due to inadequate collagen turnover.";
 RL Cell 99:81-92(1999).
 CC -!- FUNCTION: Endopeptidase that degrades various components of the
 CC extracellular matrix, such as collagen. Activates progelatinase A.
 CC Essential for pericellular collagenolysis and modeling of skeletal
 CC and extracellular connective tissues during development.
 CC -!- CATALYTIC ACTIVITY: Endopeptidase activity. Activates
 CC progelatinase A by cleavage of the propeptide at 37-Aen-Ileu-38.
 CC Other bonds hydrolyzed include 35-Gly-Ile-36 in the propeptide
 CC of collagenase 3, and 341-Asn-Phe-342, 441-Asp-Ileu-442 and
 CC 354-Gln-Thr-355 in the aggrecan interglobular domain.
 CC -!- COFACTOR: Binds 1 zinc ion per subunit, calcium (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (potential). heart,
 CC -!- TISSUE SPECIFICITY: Highly expressed in placenta, kidney, heart,
 CC lung, embryonic skeletal and periskeletal tissues.
 CC -!- DEVELOPMENTAL STAGE: Not detected before day 10.5. At day 12.5,
 CC prominently expressed in large arteries and the umbilical

arteries, expressed at lower levels in the myocardium, craniofacial mesenchyme, nasal epithelium and liver capsule. At days 14.5 and 17.5, expressed in the musculoskeletal system, and ossification areas, with continued expression in the arterial tunica media.

-!- PTM: The precursor is cleaved by a furin endopeptidase (By

similarity). Belongs to peptidase family M10A.

-!- SIMILARITY: Contains 1 hemopexin-like domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; X83536; CAA58520.2; -.
EMBL; AF022432; AAB86602.1; JOINED.
EMBL; AF022424; AAB86602.1; JOINED.
EMBL; AF022425; AAB86602.1; JOINED.
EMBL; AF022426; AAB86602.1; JOINED.
EMBL; AF022427; AAB86602.1; JOINED.
EMBL; AF022428; AAB86602.1; JOINED.
EMBL; AF022429; AAB86602.1; JOINED.
EMBL; AF022430; AAB86602.1; JOINED.
EMBL; AF022431; AAB86602.1; JOINED.
EMBL; U54984; AAB51753.1; -.
HSP; P08254; IHFS.
MEROPS; M10.014; -.
MGD; MGI:101900; Mmp14.
InterPro; IPR000585; Hemopexin.
InterPro; IPR001818; Pept_M10A_M12B.
InterPro; IPR006025; Pept_M_Zn_BS.
InterPro; IPR006026; Peptidase_M.
Pfam; PF00413; Hemopexin; 4.
Pfam; PF00415; Peptidase_M10; 1.
Pfam; PF03933; Peptidase_M10_N; 1.
PRINTS; PR00138; MATRIXIN.
SMART; SM00120; HX; 4.
SMART; SM00235; ZnMC; 1.
PROSITE; PS00024; HEMOPEXIN; 1.
PROSITE; PS00142; ZINC_PROTEASE; 1.
PROSITE; PS00546; CYSTEINE_SWITCH; 1.
Hydrolase; Metalloprotease; Zinc; Zincogen; Calcium; Signal;
Transmembrane; Collagen degradation; Extracellular matrix.
SIGNAL 1 20
FT PROPEP 21 111
FT CHAIN 112 582
FT DOMAIN 112 541
FT TRANSMEM 542 562
FT DOMAIN 563 582
FT SITE 316 511
FT SITE 93 93
FT METAL 239 239
FT ACT SITE 240 240
FT METAL 243 243
FT METAL 249 249
FT DISULFID 319 508
FT CONFLICT 133 133
FT CONFLICT 255 255
FT CONFLICT 258 258
FT CONFLICT 341 341
FT CONFLICT 346 346
FT CONFLICT 378 378
FT CONFLICT 390 391
FT CONFLICT 400 401
FT CONFLICT 407 407
FT CONFLICT 412 412
FT CONFLICT 417 417
FT CONFLICT 512 512
SEQUENCE 582 AA; 65935 MW; 3AB355158D4DD175 CRO64;

Query Match 80.5%; Score 33; DB 1; Length 582;
Best Local Similarity 87.5%; Pred. No. 23;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 ARSLILPL 9
DB 8 SRSLILPL 15

RESULT 6
MM14_RAT STANDARD; PRT; 582 AA.
ID MM14_RAT
AC Q10739; (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DT 10-OCT-2003 (Rel. 42, Matrix metalloproteinase-14 precursor (EC 3.4.24.80) (MMP-14) (MTMMP1))
DE Membrane-type matrix metalloproteinase 1 (MT-MMP 1) (MTMMP1)
DE Membrane-type-1 matrix metalloproteinase (MT1-MMP) (MTMMP) (MT-MMP)
DE MMP14 OR MTMMP.
GN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95224014; PubMed=7708715;
RA Okada A., Belloq J.P., Rouyer N., Chenard M.P., Rio M.C., Chambon P., Bassot P.;
RT "Membrane-type matrix metalloproteinase (MT-MMP) gene is expressed in stromal cells of human colon, breast, and head and neck carcinomas";
RL Proc. Natl. Acad. Sci. U.S.A. 92:2730-2734(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Cossins J., Clements J., Catlin G., Wells G.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Seems to specifically activate progelatinase A. May thus trigger invasion by tumor cells by activating progelatinase A on the tumor cell surface.
CC -!- CATALYTIC ACTIVITY: Endopeptidase activity. Activates progelatinase A by cleavage of the propeptide at 37-Asn-Leu-38. Other bonds hydrolyzed include 35-Gly-Leu-36 in the propeptide of collagenase 3, and 341-Asn-Phe-342, 441-Asp-Leu-442 and 354-Gln-Thr-355 in the aggrecan interglobular domain.
CC -!- COFACTOR: Binds 1 zinc ion per subunit, calcium (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- SIMILARITY: Belongs to peptidase family M10A.
CC -!- SIMILARITY: Contains 1 hemopexin-like domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; X83537; CAA58521.1; -.
EMBL; X91785; CAA62897.1; -.
PIR; I84471; I84471.
HSP; P08254; IHFS.
MEROPS; M10.014; -.
InterPro; IPR000585; Hemopexin.
InterPro; IPR001818; Pept_M10A_M12B.
InterPro; IPR006025; Pept_M_Zn_BS.
InterPro; IPR006026; Peptidase_M.
Pfam; PF00045; hemopexin; 4.
Pfam; PF00413; Peptidase_M10; 1.
Pfam; PF03933; Peptidase_M10_N; 1.
PRINTS; PR00138; MATRIXIN.
SMART; SM00120; HX; 4.

DR SMART, SM00235; ZnMc; 1.
 DR PROSITE; PS00024; HEMOPEXIN; 1.
 DR PROSITE; PS00142; ZINC PROTEASE; 1.
 DR PROSITE; PS00546; CYSTEINE SWITCH; 1.
 KW Hydrolase; Metalloprotease; Zinc; Zymogen; Calcium; Signal;
 KW Transmembrane.
 FT SIGNAL 1 20 POTENTIAL.
 FT PROPEP 21 111 ACTIVATION PEPTIDE.
 FT CHAIN 112 582 MATRIX METALLOPROTEINASE-14.
 FT DOMAIN 112 541 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 542 562 POTENTIAL.
 FT DOMAIN 563 582 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 316 511 HEMOPEXIN-LIKE.
 FT SITE 93 93 CYSTEINE SWITCH (POTENTIAL).
 FT METAL 239 239 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 240 240 BY SIMILARITY.
 FT METAL 243 243 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 249 249 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 319 508 BY SIMILARITY.
 FT DISULFID 68 68 I -> M (IN REF. 2).
 FT CONFLICT 255 255 D -> A (IN REF. 2).
 FT CONFLICT 255 255 D -> A (IN REF. 2).
 SQ SEQUENCE 582 AA; 66106 MW; 8840FDD9999CA80C CRC64;
 Query Match 80.5%; Score 33; DB 1; Length 582;
 Best Local Similarity 87.5%; Pred. No. 23;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 ARSLLLPL 9
 Db 8 SRSLLLPL 15
 RESULT 7
 NM14_PIG STANDARD; PRT; 580 AA.
 AC OX2T90;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Matrix metalloproteinase-14 precursor (BC 3.4.24.80) (MMP-14)
 DE (Membrane-type matrix metalloproteinase 1) (MT-MMP 1) (MTMMP1)
 DE (Membrane-type-1 matrix metalloproteinase) (MT1-MMP) (MTMMP).
 GN MMP14.
 OS Sus scrofa (Pig)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99095929; PubMed=9881602;
 RA Caron C., Xue J., Bartlett J.D.,
 RT "Expression and localization of membrane type 1 matrix
 RL metalloproteinase in tooth tissues."
 RL Matrix Biol. 17:501-511(1998).
 CC -!- FUNCTION: Seems to specifically activate progelatinase A. May thus
 CC trigger invasion by tumor cells by activating progelatinase A on
 CC the tumor cell surface (By similarity). May play a role in the
 CC biomineralization of enamel and dentin.
 CC -!- CATALYTIC ACTIVITY: Endopeptidase activity. Activates
 CC progelatinase A by cleavage of the propeptide at 37-Asn--Leu-38.
 CC Other bonds hydrolyzed include 35-Gly--Ile-36 in the propeptide
 CC of collagenase 3, and 341-Asn--Phe-342, 441-Asp--Leu-442 and
 CC 354-Gln--Thr-355 in the aggrecan interglobular domain.
 CC -!- COFACTOR: Binds 1 zinc ion per subunit, calcium (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -!- TISSUE SPECIFICITY: Highly expressed in developing tooth tissues.
 CC -!- SIMILARITY: Belongs to peptidase family M10A.
 CC -!- SIMILARITY: Contains 1 hemopexin-like domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@sb-sib.ch).
 CC -----
 DR EMBL; AF067419; AAD38324.1; -
 DR HSSP; P08254; 1HFS.
 DR MEROPS; M10.014; -
 DR InterPro; IPR000585; Hemopexin.
 DR InterPro; IPR001818; Pept_M10A_M12B.
 DR InterPro; IPR006025; Pept_M_2n_BS.
 DR InterPro; IPR006026; Peptidase_M.
 DR Pfam; PF00045; Hemopexin; 4.
 DR Pfam; PF00413; Peptidase_M10; 1.
 DR PRINTS; PR003933; Peptidase_M10_N; 1.
 DR SMART; SM00138; MATRIXIN.
 DR SMART; SM00120; HX; 4.
 DR SMART; SM00235; ZnMc; 1.
 DR PROSITE; PS00024; HEMOPEXIN; 1.
 DR PROSITE; PS00142; ZINC PROTEASE; 1.
 DR PROSITE; PS00546; CYSTEINE SWITCH; 1.
 KW Hydrolase; Metalloprotease; Zinc; Zymogen; Calcium; Signal;
 KW Transmembrane.
 FT SIGNAL 1 28 POTENTIAL.
 FT PROPEP 29 109 ACTIVATION PEPTIDE.
 FT CHAIN 110 580 MATRIX METALLOPROTEINASE-14.
 FT DOMAIN 110 539 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 540 560 POTENTIAL.
 FT DOMAIN 561 580 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 314 509 HEMOPEXIN-LIKE.
 FT SITE 91 91 CYSTEINE SWITCH (POTENTIAL).
 FT METAL 237 237 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 238 238 BY SIMILARITY.
 FT METAL 241 241 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 247 247 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 317 506 BY SIMILARITY.
 SQ SEQUENCE 580 AA; 65934 MW; B7B2C2C569A96CAC CRC64;
 Query Match 78.0%; Score 32; DB 1; Length 580;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 RSLLLPL 9
 Db 9 RSLLLPL 15
 RESULT 8
 SNXD_HUMAN STANDARD; PRT; 968 AA.
 ID SNXD_HUMAN
 AC Q9Y5W8; O94821; Q8WVZ2; Q8WXH8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Sorting nexin 13 (RGS domain- and PHOX domain-containing protein)
 DE (RGS-PX).
 GN SNX13 OR KIAA0713.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX TISSUE=Brain;
 RX MEDLINE=99087487; PubMed=9872452;
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Miyajima N.,
 RA Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XI.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro."
 RL DNA Res. 5:277-286(1998).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RA Zheng B., Farquhar M.G.;

RT "RGS-PX1, a GAP for Gαphas and a sorting nexin in vesicular
RT trafficking";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 607-968 FROM N.A.
RX MEDLINE=21378165; PubMed=1485546;
RA Teasdale R.D., Loc D., Houghton F., Karlsson L., Gleeson P.A.;
RT "A large family of endosome-localized proteins related to sorting
RT nexin 1";
RL Biochem. J. 358:7-16(2001).
RN [4]
RP SEQUENCE OF 651-968 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Krausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klauer R.D., Collins F.S., Wagner L., Shermen C.M., Schuler G.D.,
RA Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay J.J., Rulyk S.W.,
RA Villalon D.K., Muny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smalius D.E.,
RA Schnerch A., Schein J.E., Jones S.M., Marra M.A.;
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: May be involved in several stages of intracellular
CC trafficking. Act as a GAP for Gαphas.
CC -1- ALTERNATIVE PRODUCTS:
CC -1- ALTERNATIVE SPLICING: Named isoforms=2;
CC Name=1;
CC IsoId=Q9V5W8-1; Sequences=Displayed;
CC Name=2;
CC IsoId=Q9V5W8-2; Sequence=VSP_006192;
CC Note=No experimental confirmation available;
CC -1- SIMILARITY: Belongs to the sorting nexin family.
CC -1- SIMILARITY: Contains 1 RGS domain.
CC -1- SIMILARITY: Contains 1 phox homology (PX) domain.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 887.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB018256; BAA34433.1; ALT_FRAME.
DR EMBL; AF420470; AAL37728.1; -
DR EMBL; AF121862; AAD27835.1; -
DR EMBL; BC022060; AAH22060.1; ALT_INIT.
DR EMBL; HGNC:21335; SNX13.
DR MIM; 606589; -
DR InterPro; IPR001693; PX.
DR InterPro; IPR003114; PX_assoc.
DR InterPro; IPR000342; Regl_Gproteins.
DR Pfam; PF00787; PX; 1.
DR Pfam; PF02194; PXA; 1.
DR Pfam; PF00615; RGS; 1.
DR SMART; SM00312; PX; 1.
DR SMART; SM00313; PXA; 1.
DR SMART; SM00315; RGS; 1.

DR PROSITE; P850195; PX; 1.
DR PROSITE; P850132; RGS; 1.
KW Transport; protein transport; signal transduction inhibitor;
KW Alternative splicing.
FT DOMAIN 57 284 PXA.
FT DOMAIN 373 496 RGS.
FT DOMAIN 570 691 PX.
FT VARSPLIC 569 579 Missing (in isoform 2).
FT /FTID=VSP_006192.
FT CONFLICT 638 638 G -> E (IN REF. 1).
SQ SEQUENCE 968 AA; 112188 MW; 532994AE0B347FB5 CRC64;

Query Match 78.0%; Score 32; DB 1; Length 968;
Best Local Similarity 66.7%; Pred. No. 62;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
Db 254 LARGILLPL 262

RESULT 9

YATE SCHPO STANDARD; PRT; 1628 AA.
ID YATE SCHPO STANDARD; PRT; 1628 AA.
AC Q09779; O13884;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein C1D4.14 in chromosome I.
GN SPAC1D4.14 OR SPAC222F3.14C
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21384801; PubMed=11859360;
RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Collins K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA Jones K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymoprez B.,
RA Widdjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle S.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leleau V., Motier S.,
RA Galibert F., Aves S., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe";
RL Nature 415:871-880(2002).
CC -1- SIMILARITY: Belongs to the THOC2 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC -----


```

CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; Z69239; CAA93223.1; -.
CC RA BURLAND W., LIU S.-R., PLUNKETT G. III, MAYHEW G.F., ROSE D.J.,
CC RA BURLAND V., KODOYIANNI V., SCHWARTZ D.C., BLATTNER F.R.,
CC RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
CC and CT18."
CC RL J. Bacteriol. 185:2330-2337(2003).
CC CC -!- FUNCTION: SULA IS A COMPONENT OF THE SOS SYSTEM AND AN INHIBITOR
CC OF CELL DIVISION. ACCUMULATION OF SULA CAUSES RAPID CESSATION OF
CC CELL DIVISION AND THE APPEARANCE OF LONG, NON-SEPTATE FILAMENTS.
CC THE EXPRESSION OF SULA IS REPPRESSED BY LEXA PROTEIN. FTSZ SEEMS
CC TO BE THE TARGET OF SULA.
CC CC -!- SUBCELLULAR LOCATION: Inner membrane.
CC CC -!- MISCELLANEOUS: THE N-TERMINUS OF SULA PROTEIN MIGHT BE INVOLVED IN
CC RECOGNIZING THE CELL DIVISION APPARATUS.
CC CC -!- SIMILARITY: THE C-TERMINUS SHOWS SIMILARITY TO THE N PROTEIN OF
CC PHAGE LAMEDA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; M16324; AAA27230.1; -.
CC DR EMBL; AE008746; AAL20004.1; -.
CC DR EMBL; AL627269; CAD08197.1; -.
CC DR EMBL; AE016840; AAO69467.1; -.
CC DR PIR; B29016; B29016.
CC DR StyGene; SG10386; sula.
CC DR InterPro; IPR004596; Sula.
CC DR Pfam; PF03846; Sula; 1.
CC DR TIGRFAMs; TIGR00623; sula; 1.
CC DR Cell division; Septation; SOS response; Inner membrane;
CC KW Complete proteome.
CC FT DOMAIN 31 149 CONSERVED REGION.
CC FT DOMAIN 160 169 LON PROTEIN BINDING SITE (PROBABLE).
CC FT CONFLICT 102 102 S -> T (IN REF. 1).
CC FT CONFLICT 108 108 R -> A (IN REF. 1).
CC FT CONFLICT 151 151 H -> L (IN REF. 1).
CC SQ SEQUENCE 169 AA; 19013 MW; 3848A73595ED176 CRC64;
CC
CC Query Match 75.6%; Score 31; DB 1; Length 169;
CC Best Local Similarity 77.8%; Pred. No. 19;
CC Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
CC
CC QY 1 MARSLLLPL 9
CC Db 45 MAQLLLLPL 53
CC
CC RESULT 11
CC NU2M EQUAS STANDARD; PRT; 346 AA.
CC ID NU2M EQUAS AC P92476;
CC DT 15-JUL-1998 (Rel. 36, Created)
CC DT 15-JUL-1998 (Rel. 36, Last sequence update)
CC DE NADH-ubiquinone oxidoreductase chain 2 [EC 1.6.5.3].
CC GN MTND2 OR ND2 OR NADH2.
CC OS Equus asinus (Donkey).
CC OG Mitochondrion.
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
CC OX NCBI_TaxID=9793;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Kidney;
CC RX MEDLINE=97032591; PubMed=8875857;
CC RA Xu X., Gullberg A., Arnason U.;
CC RT "The complete mitochondrial DNA (mtDNA) of the donkey and mcdNA

```


comparisons among four closely related mammalian species-pairs.":
 J. Mol. Evol. 43:438-463(1996).
 -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial inner membrane.
 -!- SIMILARITY: Belongs to the complex I subunit 2 family.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

 EMBL; X97337; CRA66015.1; -.
 PIR; T11364.
 InterPro; IPR003917; NADH_oxred2.
 InterPro; IPR001750; Oxidored_q1.
 Pfam; PF00361; Oxidored_q1; 1.
 PRINTS; PR01436; NADH_OXRED2.
 Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
 SEQUENCE 346 AA; 38870 MW; 021D6D976DB564DB CRC64;

 Query Match 75.6%; Score 31; DB 1; Length 346;
 Best Local Similarity 77.8%; Pred. No. 37;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

 QY 1 MARSLLPL 9
 :| |||||
 Db 330 IASSLLPL 338

 RESULT 12
 CIND RAT STANDARD; PRT; 2005 AA.
 AC P04775;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Sodium channel protein type II alpha subunit (Sodium channel protein, brain II alpha subunit).
 GN SCN2A.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86146901; PubMed=3754035;
 RA Noda M., Ikeda T., Kayano T., Suzuki H., Takeshima H., Kurasaki M., Takahashi H., Numa S.;
 RA "Existence of distinct sodium channel messenger RNAs in rat brain."; Nature 320:188-192(1986).
 RL Nature 320:188-192(1986).
 CC -!- FUNCTION: Mediates the voltage-dependent sodium ion permeability of excitable membranes. Assuming opened or closed conformations in response to the voltage difference across the membrane, the protein forms a sodium-selective channel through which Na(+) ions may pass in accordance with their electrochemical gradient.
 CC -!- SUBUNIT: The sodium channel consists of a large polypeptide and 2-3 smaller ones. This sequence represents a large polypeptide.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- DOMAIN: The sequence contains 4 internal repeats, each with 5 hydrophobic segments (S1,S2,S3,S5,S6) and one positively charged segment (S4). Segments S4 are probably the voltage-sensors and are characterized by a series of positively charged amino acids at every third position.
 CC -!- SIMILARITY: Belongs to the sodium channel family.
 CC -!- SIMILARITY: Contains 1 IQ domain.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

 EMBL; X03639; CAA27287.1; -.
 PDB; 1BYI; 29-OCT-99.
 InterPro; IPR001682; Ca/Na_pore.
 InterPro; IPR002111; Cat_channel_Trlp.
 InterPro; IPR005821; Ion_trans.
 InterPro; IPR000048; IQ_region.
 InterPro; IPR005820; M+channel_nlg.
 InterPro; IPR001696; Na_channel.
 InterPro; IPR003915; PKD_2.
 Pfam; PF00520; ion_trans; 4.
 PRINTS; PF00612; IQ; 1.
 PRINTS; PR00170; NACHANNEL.
 PRINTS; PR01433; POLYCYSTIN2.
 SMART; SM00015; IQ; 1.
 PROSITE; PS50096; IQ; 1.
 Ionic channel; Transmembrane; Ion transport; Voltage-gated channel; Glycoprotein; Repeat; Multigene family; 3D-structure; Sodium channel.
 REPEAT 111 456
 REPEAT 741 1013
 FT REPEAT 1190 1504
 FT REPEAT 1513 1811
 FT TRANSMEM 125 148
 FT TRANSMEM 157 176
 FT TRANSMEM 190 208
 FT TRANSMEM 215 234
 FT TRANSMEM 251 274
 FT TRANSMEM 402 427
 FT TRANSMEM 754 778
 FT TRANSMEM 790 813
 FT TRANSMEM 822 841
 FT TRANSMEM 848 867
 FT TRANSMEM 884 904
 FT TRANSMEM 958 983
 FT TRANSMEM 1204 1227
 FT TRANSMEM 1241 1266
 FT TRANSMEM 1273 1294
 FT TRANSMEM 1299 1320
 FT TRANSMEM 1340 1367
 FT TRANSMEM 1447 1473
 FT TRANSMEM 1527 1550
 FT TRANSMEM 1562 1585
 FT TRANSMEM 1592 1615
 FT TRANSMEM 1626 1647
 FT TRANSMEM 1663 1685
 FT TRANSMEM 1752 1776
 FT TRANSMEM 1905 1934
 FT DOMAIN
 FT CARBOHYD 212 212
 FT CARBOHYD 285 285
 FT CARBOHYD 291 291
 FT CARBOHYD 297 297
 FT CARBOHYD 303 303
 FT CARBOHYD 308 308
 FT CARBOHYD 340 340
 FT CARBOHYD 604 604
 FT CARBOHYD 624 624
 FT CARBOHYD 883 883
 FT CARBOHYD 1055 1055
 FT CARBOHYD 1072 1072
 FT CARBOHYD 1136 1136
 FT CARBOHYD 1368 1368
 FT CARBOHYD 1382 1382
 FT CARBOHYD 1393 1393
 SQ SEQUENCE 2005 AA; 227872 MW; 861BE583D79F8324 CRC64;

 Query Match 75.6%; Score 31; DB 1; Length 2005;
 Best Local Similarity 75.0%; Pred. No. 2; 1e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 MARSLLLP 8
Db 1 MARSVLVP 8

RESULT 13
CYTC_RABIT STANDARD; PRT; 148 AA.
AC O97852;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cystatin C precursor.
GN OS
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=Japanese white; TISSUE=Bone;
RC MEDLINE=98424349; PubMed=9753427;
RX Kabori M., Ikeda Y., Nara H., Kato M., Kumegawa M., Nojima H.,
RA Kawashima H.;
RA "Large scale isolation of osteoclast-specific genes by an improved
RT method involving the preparation of a subtracted cDNA library.";
RL Genes Cells 3:49-475(1998).
CC -!- FUNCTION: This is a thiol proteinase inhibitor.
CC -!- SIMILARITY: Belongs to the cystatin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB009342; BAA75921.1; -
DR HSPB; P01034; 1096.
DR InterPro; IPR000010; Cystatin.
DR SMART; PF00031; cystatin; 1.
DR Pfam; SMO0043; Cys; 1.
DR PROSITE; PS00287; CYSTATIN; FALSE NEG.
KW Thiol protease inhibitor; Signal.
FT SIGNAL 1 28
FT CHAIN 29 148
FT ACT_SITE 39 39
FT SITE 83 87
FT DISULFID 101 111
FT DISULFID 125 145
FT SEQUENCE 148 AA; 16346 MW; 1523C8311695B9A CRC64;

Query Match 73.2%; Score 30; DB 1; Length 148;
Best Local Similarity 77.8%; Pred. No. 27;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 MARSLLLP 9
Db 1 MARSGLVPL 9

RESULT 14
CORZ_DROME STANDARD; PRT; 154 AA.
AC Q26377; Q9VFK7;
DT 15-JUL-1998 (Rel. 36, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Crz precursor [Contains: Corazonin; Corazonin-precursor-related
DE Peptide (CRP)].
GN OS
OS Drosophila melanogaster (Fruit fly).

```

```

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=Berkley;
RC MEDLINE=2019606; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Buesan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Morkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J., Yao Q.A.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RN Science 287:2185-2195(2000).
RN [2]
RN SEQUENCE OF 1-82 FROM N.A.
RP MEDLINE=95032110; PubMed=7945373;
RX Veenstra J.A.;
RT "Isolation and structure of the Drosophila corazonin gene.";
RL Biochem. Biophys. Res. Commun. 204:292-296(1994).
RN [3]
RN SEQUENCE OF 20-30, AND AMIDATION.
RC TISSUE=Larva;
RX MEDLINE=22287343; PubMed=12171930;
RA Baggerman G., Cerstiaens A., De Loof A., Schoofs L.;
RT "Peptidomics of the larval Drosophila melanogaster central nervous
RT system.";
RL J. Biol. Chem. 277:40368-40374(2002).
CC -!- FUNCTION: Cardioactive peptide. Corazonin is probably involved
CC in the physiological regulation of the heart beat.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE003704; AAP55046.1; -
DR

```

Db 102 IAKSLLLP 109
Search completed: March 1, 2004, 17:30:01
Job time : 7 secs

DR EMBL; S74038; AAB32283.1; ALT_SEQ.
DR FlyBase; FBgn0013767; Crz.
KW Neuropeptide; Amidation; Cleavage on pair of basic residues; Signal;
KW Pyrrolidone carboxylic acid.
FT SIGNAL 1 19
FT PEPTIDE 20 30 CORAZONIN.
FT PEPTIDE 34 67 CORAZONIN-PRECURSOR-RELATED PEPTIDE.
FT PROPEP 68 154
FT MOD RES 20 20 PYRROLIDONE CARBOXYLIC ACID.
FT MOD RES 30 30 AMIDATION (G-31 PROVIDE AMIDE GROUP).
FT CONFLICT 81 81 L -> Q (IN REF. 2).
SQ SEQUENCE 154 AA; 17154 MW; 06046282A1C61A2F CRC64;

Query Match 73.2%; Score 30; DB 1; Length 154;
Best Local Similarity 77.8%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MARSLLLP 9
DB 1 MLRULLLP 9

RESULT 15
YHJC_ECOLI
ID YHJC_ECOLI STANDARD; PRT; 299 AA.
AC P37641;
DT 01-OCT-1994 (Rel. 30, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Putative HTH-type transcriptional regulator YHJC.
GN YHJC OR B3521.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=94316500; PubMed=8041620;
RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner P.R.;
RT "Analysis of the Escherichia coli genome. V. DNA sequence of the
RT region from 76,0 to 81.5 minutes."
RL Nucleic Acids Res. 22:2576-2586(1994).
CC -!- SIMILARITY: Contains 1 HTH lySR-type DNA-binding domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

CC EMBL; U00039; AAB18497.1; ALT_INIT.
DR EMBL; A3000428; AAC76546.1; ALT_INIT.
DR EcoGene; EGI2247; YHJC.
DR InterPro; IPR000847; HTH_LySR.
DR InterPro; IPR005119; LySR_subst.
DR Pfam; PF00126; HTH_1; 1.
DR Pfam; PF03466; LySR_substrate; 1.
DR PROSITE; PS50931; HTH_LySR; 1.
KW Hypothetical protein; Transcription regulation; DNA-binding;
KW Complete proteome.
FT DOMAIN 1 59 HTH_LySR-TYPE.
FT DNA_BIND 19 38 H-T-H MOTIF (BY SIMILARITY).
SQ SEQUENCE 299 AA; 33329 MW; 22D0EC5994C8C8C1 CRC64;

Query Match 73.2%; Score 30; DB 1; Length 299;
Best Local Similarity 75.0%; Pred. No. 53;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLP 8
DB 1 MLRULLLP 8

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:20:41 ; Search time 30.3333 Seconds
(without alignments)
93.615 Million cell updates/sec

Title: US-09-905-083-86

Perfect score: 41

Sequence: 1 MARSLLLPL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organalle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	253	4 Q8N5N9	Q8N5n9 homo sapien
2	35	85.4	296	4 Q9NZL1	Q9nzl1 homo sapien
3	35	85.4	468	4 Q9NZL2	Q9nzl2 homo sapien
4	35	85.4	1135	10 Q8W4W9	Q8w4w9 arabidopsis
5	35	85.4	1192	10 Q9S7T0	Q9s7t0 arabidopsis
6	34	82.9	129	10 Q8H584	Q8h584 oryza sativ
7	34	82.9	348	16 Q7U4K2	Q7u4k2 synechococc
8	33	80.5	272	10 Q8LC74	Q8lc74 arabidopsis
9	33	80.5	272	10 Q9PNC4	Q9pnc4 arabidopsis
10	33	80.5	295	10 Q8W2V5	Q8w2v5 oryza sativ
11	33	80.5	295	10 Q7XFE9	Q7xfe9 oryza sativ
12	33	80.5	368	10 Q8SAT6	Q8sat6 oryza sativ
13	33	80.5	397	16 Q9AA49	Q9aa49 caulobacter
14	33	80.5	424	16 Q827F2	Q827f2 streptomyce
15	33	80.5	526	10 Q9ZRH9	Q9zrh9 oryza sativ
16	33	80.5	571	10 Q9SYC9	Q9syc9 arabidopsis

17	33	80.5	582	11 Q8BTX2	Q8bt2 mus musculus
18	33	80.5	809	5 Q8SS23	Q8ss23 encephalito
19	33	80.5	934	13 Q9DER4	Q9der4 gallus gall
20	33	80.5	2998	5 Q8MXK6	Q8mxk6 leishmania
21	32	78.0	81	16 Q92GT3	Q92gt3 rickettsia
22	32	78.0	126	16 Q9KRE6	Q9kre6 vibrio chol
23	32	78.0	149	16 Q8KAK3	Q8kak3 chlorobium
24	32	78.0	226	4 Q8NH11	Q8nh11 homo sapien
25	32	78.0	227	4 Q9UKJ0	Q9ukj0 homo sapien
26	32	78.0	247	2 Q9KWM3	Q9kwm3 staphylococ
27	32	78.0	271	4 Q9HBS0	Q9hbs0 homo sapien
28	32	78.0	274	16 Q8EAX3	Q8eax3 shewanella
29	32	78.0	303	4 Q9UKJ1	Q9ukj1 homo sapien
30	32	78.0	306	6 Q9SKJ4	Q9skj4 macaca fasc
31	32	78.0	334	16 Q988I7	Q988i7 rhizobium 1
32	32	78.0	363	16 Q8PJF1	Q8pjf1 xanthomonas
33	32	78.0	425	16 Q88M7	Q88m7 pseudomonas
34	32	78.0	432	16 Q9HV31	Q9hvb1 pseudomonas
35	32	78.0	444	16 Q88Q03	Q88q03 pseudomonas
36	32	78.0	506	16 Q9Z5I3	Q9z5i3 mycobacteri
37	32	78.0	544	16 Q82PC2	Q82pc2 streptomyce
38	32	78.0	654	6 Q95LV1	Q95lv1 macaca fasc
39	32	78.0	654	6 Q95LY1	Q95ly1 macaca fasc
40	32	78.0	665	16 Q834W2	Q834w2 enterococcu
41	32	78.0	921	10 Q9SDQ4	Q9sdq4 arabidopsis
42	32	78.0	921	10 Q9ZV68	Q9zv68 arabidopsis
43	32	78.0	957	4 Q8XCA4	Q8xca4 homo sapien
44	31	75.6	138	11 Q8Q463	Q8q463 cricetus
45	31	75.6	147	2 Q44395	Q44395 agrobacteri

ALIGNMENTS

RESULT 1

Q8N5N9 PRELIMINARY; PRT; 253 AA.
AC Q8N5N9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Kallikrein 7 (chymotryptic, stratum corneum).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strauberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; BC032005; AAH32005.1; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00202; TRYPSIN_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 253 AA; 2D68B6A41B22A668 CRC64;

Query March 100.0%; Score 41; DB 4; Length 253;

Best Local Similarity 100.0%; Pred. No. 1.6;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
| | | | |
DB 1 MARSLLLPL 9

RESULT 2
Q9NZL1 PRELIMINARY; PRT; 396 AA.
AC Q9NZL1; 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Aspartyl protease.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20422477; PubMed=10965118;
RA Solans A., Estivill X., de La Luna S.;
RT "A new aspartyl protease on 21q22.3, BACE2, is highly similar to
RT Alzheimer's amyloid precursor protein beta-secretase.";
RC Cytogenet. Cell Genet. 89:177-184(2000).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
DR EMBL; AF188277; AAF35836.1; --
DR HSSP; P00797; 2REN.
DR GO; GO:0004194; F:pepsin A activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001969; Asparticase AS.
DR InterPro; IPR001461; Peptidase A1.
DR InterPro; IPR009007; Pept_A_acid.
DR Pfam; PF00026; asp; 1
DR PRINTS; P00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
DR Aspartyl protease; Hydrolase; Protease.
KW ASPARTYL PROTEASE; HYDROLASE; PROTEASE.
SQ SEQUENCE 396 AA; 43013 MW; 5023A7AF391CEAC9 CRC64;

Query Match 85.4%; Score 35; DB 4; Length 396;
Best Local Similarity 77.8%; Pred. No. 46;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
| | | | |
DB 4 LARALLPL 12

RESULT 3
Q9NZL2 PRELIMINARY; PRT; 468 AA.
AC Q9NZL2; 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Aspartyl protease.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20422477; PubMed=10965118;
RA Solans A., Estivill X., de La Luna S.;
RT "A new aspartyl protease on 21q22.3, BACE2, is highly similar to
RT Alzheimer's amyloid precursor protein beta-secretase.";
RC Cytogenet. Cell Genet. 89:177-184(2000).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
DR EMBL; AF188276; AAF35835.1; --
DR HSSP; P00797; 2REN.
DR GO; GO:0004194; F:pepsin A activity; IEA.

DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001969; Asparticase AS.
DR InterPro; IPR001461; Peptidase A1.
DR InterPro; IPR009007; Pept_A_acid.
DR Pfam; PF00026; asp; 1.
DR PRINTS; P00792; PEPsin.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
DR Aspartyl protease; Hydrolase; Protease.
KW ASPARTYL PROTEASE; HYDROLASE; PROTEASE.
SQ SEQUENCE 468 AA; 50324 MW; 71780920126A0142 CRC64;

Query Match 85.4%; Score 35; DB 4; Length 468;
Best Local Similarity 77.8%; Pred. No. 54;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
| | | | |
DB 4 LARALLPL 12

RESULT 4
Q84W49 PRELIMINARY; PRT; 1135 AA.
AC Q84W49; 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
GN AT3G01780.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,
RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Wong C., Wu H.C.,
RA Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J., Narusaka M.,
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,
RA Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Arabidopsis Full Length cDNA Clones.";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT004227; AAO42242.1; --
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR002048; EF-hand.
DR PROSITE; PS00018; EF_HAND; 1.
KW Hypothetical protein.
FT NON TER 1135 1135
SQ SEQUENCE 1135 AA; 126119 MW; C5FDDC178D1E2D96 CRC64;

Query Match 85.4%; Score 35; DB 10; Length 1135;
Best Local Similarity 77.8%; Pred. No. 13e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLLPL 9
| | | | |
DB 237 MARSLVLPV 245

RESULT 5
Q9S7T0 PRELIMINARY; PRT; 1192 AA.
AC Q9S7T0; 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE F28J7.11 protein.
GN F28J7.11 OR F4P13.33.
OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
 RA Roining C.M., Koo H., Fujii C.Y., Uterback T.R., Barnstead M.E.,
 RA Bowman C.D., White O., Nierman W.C., Fraser C.M.;
 RT "Arabidopsis thaliana chromosome III BAC F28J7 genomic sequence."
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
 RA Roining C.M., Koo H., Fujii C.Y., Uterback T.R., Barnstead M.E.,
 RA Bowman C.D., White O., Nierman W.C., Fraser C.M.;
 RT "Arabidopsis thaliana chromosome III BAC F4P13 genomic sequence."
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC010797; AAF03433.1; -;
 DR EMBL; AC009325; AAF01560.1; -;
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR InterPro; IPR008938; ARM.
 DR InterPro; IPR002048; EF-hand.
 DR PROSITE; PS00018; EF_HAND; 1.
 SQ SEQUENCE 1192 AA; 132863 MW; 4F67B124CBFAFF154 CRC64;

Query Match 85.4%; Score 35; DB 10; Length 1192;
 Best Local Similarity 77.8%; Pred. No. 1.3e+02;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLPL 9
 DB 237 MARSLLPLV 245
 |||||

RESULT 6
 QH584
 ID QH584 PRELIMINARY; PRT; 129 AA.
 AC QH584;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE QJ1656_E11.11 protein.
 GN QJ1656_E11.11.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RA "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
 RT clone:QJ1656 E11.11";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP003943; BAC24850.1; -;
 SQ SEQUENCE 129 AA; 13891 MW; 6334E639E1A6DEAE CRC64;

Query Match 82.9%; Score 34; DB 10; Length 129;
 Best Local Similarity 87.5%; Pred. No. 26;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLLPL 8
 DB 40 MARSLLPL 47
 |||||

RESULT 7
 Q7U4K2
 ID Q7U4K2 PRELIMINARY; PRT; 348 AA.

AC Q7U4K2;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative lopa protein.
 DE SYN2085.
 GN Synchococcus sp. (strain WH8102).
 OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.
 OX NCBI_TaxID=84588;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=22825697; PubMed=12917641;
 RA Palenik B., Brahamsha B., Larimer F.W., Land M., Hauser L., Chain P.,
 RA Lamerdin J., Regala W., Allen E.E., McCarren J., Paulsen I.,
 RA Dufresne A., Partensky F., Webb E.A., Waterbury J.;
 RL "The genome of a motile marine Synchococcus."
 RT Nature 424:1037-1042(2003).
 RL EMBL; BX569694; CAE08580.1; -;
 KW Complete proteome.
 SQ SEQUENCE 348 AA; 37531 MW; 1E1903E684081A95 CRC64;

Query Match 82.9%; Score 34; DB 16; Length 348;
 Best Local Similarity 77.8%; Pred. No. 67;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MARSLLPL 9
 DB 298 MARSLLPL 306
 |||||

RESULT 8
 Q8LC74
 ID Q8LC74 PRELIMINARY; PRT; 272 AA.
 AC Q8LC74;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Vegetative storage protein-like.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Haas B.J., Volkovsky N., Town C.D., Troukhan M., Alexandrov N.,
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
 RT "Full-length messenger RNA sequences greatly improve genome
 RT annotation."
 RL Genome Biol. 0:0-0(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RT "Full-length cDNA from Arabidopsis thaliana."
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY086753; AAM63804.1; -;
 DR InterPro; IPR005519; acid_phosphat_B.
 DR Pfam; PF03767; acid_phosphat_B; 1.
 SQ SEQUENCE 272 AA; 31041 MW; 330BB2AAD8E312A3 CRC64;

Query Match 80.5%; Score 33; DB 10; Length 272;
 Best Local Similarity 88.9%; Pred. No. 85;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MARSLLPL 9
 DB 1 MARSLLPL 9
 |||||

RESULT 9
 Q9FNC4
 ID Q9FNC4 PRELIMINARY; PRT; 272 AA.

Q9FNC4;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Putative vegetative storage protein.
 GN AT5G44020 OR MRH10.13.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 .OX NCBI_TaxID=3702;
 [1]
 SEQUENCE FROM N.A.
 RP STRAIN=Columbia;
 RC MEDLIN=98069011; PubMed=9405937;
 RA Kotani H., Nakamura Y., Sato S., Kaneko T., Asamizu E., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. II.
 RT Sequence features of the regions of 1,044,062 bp covered by thirteen
 RT physically assigned P1 clones.";
 RL DNA Res. 4:291-300(1997).
 RN [2]
 SEQUENCE FROM N.A.
 RP Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
 RA Ecker J.R., Theologis A., Davis R.W.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Pham P.K., Banh J., Dale J.M., Goldsmith A.D.,
 RA Jiang P.X., Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M.,
 RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
 RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
 RA Shinn P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,
 RA Ecker J.R., Theologis A.;
 RL "Full length cDNA of gene MRH10.13/AT5G44020 (GI:9758561).";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 SEQUENCE FROM N.A.
 RP Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
 RA Ecker J.R., Theologis A., Davis R.W.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN [5]
 SEQUENCE FROM N.A.
 RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hsuan V.W.,
 RA Lee J.M., Quach H.L., Tang C., Toriumi M., Wallender E.K., Wong C.,
 RA Wu H.C., Yu G., Yuan S., Carninci P., Chen H., Cheuk R.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J.,
 RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
 RA Shinn P., Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W.,
 RA Ecker J.R., Theologis A.;
 RL "Arabidopsis Open Reading Frame (ORF) Clones.";
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB006703; BAB09062.1; -
 DR EMBL; AY054542; RAK96733.1; -
 DR EMBL; AY045786; RAK76460.1; -
 DR EMBL; AY081695; RAM10257.1; -
 DR EMBL; AY142574; RAN13143.1; -
 DR InterPro; IPR005519; acid_phosphat_B.
 DR Pfam; PF03767; acid_phosphat_B; 1.
 SQ SEQUENCE 272 AA; 31056 MW; E5FE016FDAF0EA93 CRC64;
 Query Match 80.5%; Score 33; DB 10; Length 272;
 Best Local Similarity 88.9%; Pred. No. 85;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MARSLLLP 9
 DB 1 MARSLLLP 9
 RESULT 10
 Q8W2V5 PRELIMINARY; PRT; 295 AA.
 ID Q8W2V5; DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 AC Q8W2V5; DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN OSJNB0076H04.17.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 .OX NCBI_TaxID=4530;
 [1]
 SEQUENCE FROM N.A.
 RP STRAIN=cv. Nipponbare;
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
 RA Gansberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,
 RA Tsitrin T., Riggs F., Hsiao J., Zismann V., Blunt S., Pal G.,
 RA VanAken S.E., Utterback T.R., Feldblyum T.V., Kalb E., Quackenbush J.,
 RA Salzberg S.L., White O., Fraser C.M.;
 RT "Oryza sativa chromosome 10 BAC OSJNB0076H04 genomic sequence.";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC091093; AAL58137.1; -
 DR Gramene; Q8W2V5; -
 KW Hypothetical protein.
 SQ SEQUENCE 295 AA; 32804 MW; 6E06C01742DAF969 CRC64;
 Query Match 80.5%; Score 33; DB 10; Length 295;
 Best Local Similarity 87.5%; Pred. No. 92;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MARSLLLP 8
 DB 18 MARSLLFP 25
 RESULT 11
 Q7XFE9 PRELIMINARY; PRT; 295 AA.
 ID Q7XFE9; DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 AC Q7XFE9; DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN OSJNB0076H04.17
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 .OX NCBI_TaxID=39947;
 [1]
 SEQUENCE FROM N.A.
 RP STRAIN=cv. Nipponbare;
 RC STRAIN=cv. Nipponbare;
 RA The Rice Chromosome 10 Sequencing Consortium;
 RT "In-depth view of structure, activity, and evolution of rice
 RT chromosome 10.";
 RL Science 300:1566-1569(2003).
 RN [2]
 SEQUENCE FROM N.A.
 RP STRAIN=cv. Nipponbare;
 RC Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
 RA Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE017082; AAP53277.1; -
 KW Hypothetical protein.

SQ SEQUENCE 295 AA; 32804 MW; 6206C01742DAF969 CRC64;
Query Match 80.5%; Score 33; DB 10; Length 295;
Best Local Similarity 87.5%; Pred. No. 92;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MARSLLLP 8
DB 18 MARSLLFP 25
RESULT 12
Q8SAT6 PRELIMINARY; PRT; 368 AA.
AC Q8SAT6, 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OSJNBA0022D10.13.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
RA Sasaki C., Henry D., Oates R., Simmons J.,
RT "Rice Genomic Sequence."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC099402; AAL79349.1;
DR Gramene; Q8SAT6;
KW Hypothetical protein.
SQ SEQUENCE 368 AA; 41064 MW; BD1BF772DA4F043F CRC64;
Query Match 80.5%; Score 33; DB 10; Length 368;
Best Local Similarity 87.5%; Pred. No. 11e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MARSLLLP 8
DB 18 MARSLLFP 25
RESULT 13
Q9AA49 PRELIMINARY; PRT; 397 AA.
AC Q9AA49, 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Hypothetical protein CC0757.
GN CC0757.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=ATCC 19089 / CB15;
RC MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A.S., Stephens C., Phadke N.D., Ely B.,
RA DeSoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Koonin J.F., Smit J., Craven M.B., Kouri H., Shetty J., Berry K.,
RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005751; AA22742.1;
DR PIR; B87343; B87343.

TIGR; CC0757; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR005495; Yjgp_Yjgp.
DR Pfam; PF03739; Yjgp_Yjgp.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 397 AA; 42687 MW; AF0EB915C02ED58B CRC64;
Query Match 80.5%; Score 33; DB 16; Length 397;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 MARSLLLP 9
DB 285 LARALLPL 293
RESULT 14
Q827F2 PRELIMINARY; PRT; 424 AA.
AC Q827F2, 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Putative membrane transport protein.
GN SAV6972.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RC MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis; deducing the ability of producing secondary
RT metabolites."
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RC MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis."
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP005048; BAC74683.1; -.
KW Complete proteome.
SQ SEQUENCE 424 AA; 43493 MW; 500BA55A8D6BF80C CRC64;
Query Match 80.5%; Score 33; DB 16; Length 424;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 ARSLLLP 9
DB 290 ARALLPL 297
RESULT 15
Q2ZRH9 PRELIMINARY; PRT; 526 AA.
AC Q2ZRH9, 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Early embryogenesis protein.
GN OSE351.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Tainung 67;
 RA Tseng M.J., Wang C.S., Hsu H.R.;
 RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U25968; AAD10369.1; -.
 DR Gramene; Q9ZRH9; -.
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR006045; Cupin.
 DR InterPro; IPR006044; Seedstore_11s.
 DR Pfam; PF00190; Cupin; 2.
 DR PRINTS; PR00439; 11SGLOBULIN.
 SQ SEQUENCE 526 AA; 58770 MW; BC23COF9D1711F9B CRC64;

Query Match 80.5%; Score 33; DB 10; Length 526;
 Best Local Similarity 77.8%; Pred. No. 1.6e+02;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MARSLLPL 9
 |||||
 DB 6 MAASLLPL 14

Search completed: March 1, 2004, 17:34:54
 Job time : 32.3333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:21:46 ; Search time 11.889 Seconds
(without alignments)
39.081 Million cell updates/sec

Title: US-09-905-083-99
Perfect score: 42
Sequence: 1 QRIKASKSF 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCITUS_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	9	3	US-09-502-600-99
2	42	100.0	9	3	US-09-502-600-130
3	42	100.0	9	4	US-09-918-243-99
4	42	100.0	9	4	US-09-918-243-130
5	42	100.0	144	4	US-09-618-259-4
6	42	100.0	154	3	US-09-261-416-7
7	42	100.0	224	3	US-08-944-483-33
8	42	100.0	225	2	US-08-557-146-12
9	42	100.0	225	2	US-09-027-337-4
10	42	100.0	225	2	US-09-154-344-12
11	42	100.0	225	4	US-09-644-600-4
12	42	100.0	225	4	US-09-654-600A-4
13	42	100.0	223	2	US-08-557-146-2
14	42	100.0	223	2	US-08-824-874-3
15	42	100.0	223	2	US-09-154-344-2
16	42	100.0	223	3	US-08-930-188-2
17	42	100.0	223	3	US-09-210-084-3
18	42	100.0	223	4	US-09-764-762-3
19	42	100.0	223	5	PCT-US96-04294-2
20	36	85.7	9	3	US-09-502-600-97
21	36	85.7	9	4	US-09-318-243-97
22	31	73.8	35	4	US-09-149-476-619
23	31	73.8	250	1	US-07-914-282D-6
24	31	73.8	250	1	US-08-276-887A-6
25	31	73.8	250	5	PCT-US93-02460-6
26	31	73.8	2860	2	US-08-826-267-2
27	30	71.4	342	4	US-09-252-991A-18101

```

28 71.4 666 3 US-08-982-785A-11 Sequence 11, Appl
29 69.0 222 4 US-09-134-001C-3400 Sequence 3400, Ap
30 66.7 61 1 US-08-194-211A-3 Sequence 3, Appl
31 66.7 61 3 US-08-456-748B-3 Sequence 3, Appl
32 66.7 67 4 US-09-489-039A-10090 Sequence 10090, A
33 66.7 159 4 US-09-134-000C-4565 Sequence 4565, Ap
34 66.7 411 4 US-09-328-352-6424 Sequence 6424, Ap
35 66.7 438 4 US-09-489-039A-8527 Sequence 8527, Ap
36 66.7 446 4 US-09-328-352-7524 Sequence 7524, Ap
37 66.7 446 4 US-09-134-000C-3434 Sequence 3434, Ap
38 66.7 473 4 US-09-328-352-5279 Sequence 5279, Ap
39 66.7 500 4 US-09-442-100-14 Sequence 14, Appl
40 66.7 500 4 US-08-939-106-14 Sequence 14, Appl
41 66.7 500 4 US-09-442-102-14 Sequence 14, Appl
42 66.7 641 4 US-09-489-039A-8248 Sequence 8248, Ap
43 66.7 701 4 US-09-328-352-5717 Sequence 5717, Ap
44 66.7 823 4 US-09-107-532A-6343 Sequence 6343, Ap
45 66.7 1824 2 US-08-680-327-3 Sequence 3, Appl

```

ALIGNMENTS

RESULT 1
US-09-502-600-99
; Sequence 99, Application US/09502600A
; Patent No. 6294344
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP-C
; CURRENT FILING DATE: 2000-02-11
; PRIORITY APPLICATION NUMBER: US/09/502,600A
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 99
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 91-99 of the SCCE protein
US-09-502-600-99

Query Match Similarity 100.0%; Score 42; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 QRIKASKSF 9
   |||||
Db 1 QRIKASKSF 9

```

RESULT 2
US-09-502-600-130
; Sequence 130, Application US/09502600A
; Patent No. 6294344
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP-C
; CURRENT FILING DATE: 2000-02-11
; PRIORITY APPLICATION NUMBER: US/09/502,600A
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 130
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

OTHER INFORMATION: Residues 91-99 of the SCCE protein
US-09-502-600-130

Query Match 100.0%; Score 42; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9
DB 1 QRIKASKSF 9

RESULT 3

US-09-918-243-99
; Sequence 99, Application US/09918243
; Patent No. 6627403
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 99
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 91-99 of the SCCE protein
US-09-918-243-99

Query Match 100.0%; Score 42; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9
DB 1 QRIKASKSF 9

RESULT 4

US-09-918-243-130
; Sequence 130, Application US/09918243
; Patent No. 6627403
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 130
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 91-99 of the SCCE protein
US-09-918-243-130

Query Match 100.0%; Score 42; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9
DB 1 QRIKASKSF 9

RESULT 5

US-09-618-259-4
; Sequence 4, Application US/09618259
; Patent No. 6642013
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Underwood, Lowell J.
; TITLE OF INVENTION: No. 6642013el Extracellular Serine Protease
; FILE REFERENCE: D6020CIP2
; CURRENT APPLICATION NUMBER: US/09/618,259
; CURRENT FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: US 09/127,444
; PRIOR FILING DATE: 1998-08-21
; NUMBER OF SEQ ID NOS: 72
; SEQ ID NO 4
; LENGTH: 144
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; NAME/KEY: DOMAIN
; OTHER INFORMATION: Amino acid sequence of stratum corneum chymotryptic
; OTHER INFORMATION: enzyme (sccc) catalytic domain
US-09-618-259-4

Query Match 100.0%; Score 42; DB 4; Length 144;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9
DB 28 QRIKASKSF 36

RESULT 6

US-09-261-416-7
; Sequence 7, Application US/09261416A
; Patent No. 6291663
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Underwood, Lowell J.
; TITLE OF INVENTION: TAGD-12: A No. 6291663el Transmembrane Serine Protease
; TITLE OF INVENTION: Overexpressed in Ovarian Carcinoma
; FILE REFERENCE: D6192
; CURRENT APPLICATION NUMBER: US/09/261,416A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 7
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Serine protease catalytic domain of stratum corneum
; OTHER INFORMATION: chymotryptic enzyme (Sccc) homologous to similar
; OTHER INFORMATION: domain in TAGD-12
US-09-261-416-7

Query Match 100.0%; Score 42; DB 3; Length 154;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9
DB 37 QRIKASKSF 45

RESULT 7

US-08-944-403
; Sequence 13, Application US/089444483

Patent No. 6232456
GENERAL INFORMATION:
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: KLASS, MICHAEL R.
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STEWART, KENT D.
APPLICANT: STROUPE, STEVEN D.
TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
TITLE OF INVENTION: OF THE PROSTATE
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/944,483
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6183.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 224 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6232456e
US-08-944-483-33
Query Match 100.0%; Score 42; DB 3; Length 224;
Best Local Similarity 100.0%; Pred. No. 0.24; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRIKASKSF 9
Db 62 QRIKASKSF 70
RESULT 8
US-08-557-146-12
Sequence 12, Application US/08557146
Patent No. 5834290
GENERAL INFORMATION:
APPLICANT: Egelrud, Torbjorn
APPLICANT: Hansson, Lennart
TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
TITLE OF INVENTION: Enzyme (SCCE)
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: White & Case, Patent Department
STREET: 1155 Avenue of the Americas
CITY: New York

STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/557,146
FILING DATE: 14-DEC-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sterner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-181
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 225 amino acids
TYPE: amino acids
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: polypeptide
US-08-557-146-12
Query Match 100.0%; Score 42; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 0.25; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRIKASKSF 9
Db 63 QRIKASKSF 71
RESULT 9
US-09-027-337-4
Sequence 4, Application US/09027337B
Patent No. 5974616
GENERAL INFORMATION:
APPLICANT: O'Brien, Timothy J.
APPLICANT: Tanimoto, Hirotooshi
TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease Overexpressed in
TITLE OF INVENTION: Breast and Ovarian Carcinomas
FILE REFERENCE: D6064
CURRENT APPLICATION NUMBER: US/09/027,337B
CURRENT FILING DATE: 1998-02-20
NUMBER OF SEQ ID NOS: 13
SEQ ID NO 4
LENGTH: 225
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Serine protease catalytic domain of Scce homologous to
OTHER INFORMATION: similar domain in TADG-15
US-09-027-337-4
Query Match 100.0%; Score 42; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 0.25; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRIKASKSF 9
Db 63 QRIKASKSF 71
RESULT 10
US-09-154-344-12
Sequence 12, Application US/09154344
Patent No. 5981256
GENERAL INFORMATION:

APPLICANT: Egelrud, Torbjorn
 APPLICANT: Hanson, Lennart
 TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
 TITLE OF INVENTION: Enzyme (SCCE)
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: White & Case, Patent Department
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036-2787
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/154,344
 FILING DATE: 16-SEP-1998
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/557,146
 FILING DATE: 14-DEC-1995
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Sterner, Richard J.
 REGISTRATION NUMBER: 35,372
 REFERENCE/DOCKET NUMBER: 1103326-181
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 819-8783
 TELEFAX: (212) 354-8113
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 225 amino acids
 TYPE: amino acids
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: polypeptide
 US-09-154-344-12

Query Match 100.0%; Score 42; DB 2; Length 225;
 Best Local Similarity 100.0%; Pred. No. 0.25;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORIKASKSF 9
 DB 63 ORIKASKSF 71

RESULT 11.
 US-09-644-600-4
 ; Sequence 4, Application US/09644600
 ; Patent No. 6451500
 ; GENERAL INFORMATION:
 ; APPLICANT: O'Brien, Timothy J.
 ; APPLICANT: Tanimoto, Hirotochi
 ; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease
 ; TITLE OF INVENTION: Overexpressed in Carcinomas
 ; FILE REFERENCE: D6064CIP/D
 ; CURRENT APPLICATION NUMBER: US/09/644,600
 ; CURRENT FILING DATE: 2000-08-23
 ; PRIOR APPLICATION NUMBER: 09/421,213
 ; PRIOR FILING DATE: 1999-10-20
 ; PRIOR APPLICATION NUMBER: 09/027,337
 ; PRIOR FILING DATE: 1998-02-20
 ; NUMBER OF SEQ ID NOS: 98
 ; SEQ ID NO 4
 ; LENGTH: 225
 ; TYPE: PRI
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: SCCE

US-09-644-600-4

Query Match 100.0%; Score 42; DB 4; Length 225;
 Best Local Similarity 100.0%; Pred. No. 0.25;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORIKASKSF 9
 DB 63 ORIKASKSF 71

RESULT 12
 US-09-654-600A-4
 ; Sequence 4, Application US/09654600A
 ; Patent No. 6649741
 ; GENERAL INFORMATION:
 ; APPLICANT: O'Brien, Timothy J.
 ; APPLICANT: Tanimoto, Hirotochi
 ; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease
 ; TITLE OF INVENTION: Overexpressed in Carcinomas
 ; FILE REFERENCE: D6064CIP/D
 ; CURRENT APPLICATION NUMBER: US/09/654,600A
 ; CURRENT FILING DATE: 2000-09-01
 ; PRIOR APPLICATION NUMBER: 09/421,213
 ; PRIOR FILING DATE: 1999-10-20
 ; NUMBER OF SEQ ID NOS: 98
 ; SEQ ID NO 4
 ; LENGTH: 225
 ; TYPE: PRI
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: SCCE
 US-09-654-600A-4

Query Match 100.0%; Score 42; DB 4; Length 225;
 Best Local Similarity 100.0%; Pred. No. 0.25;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORIKASKSF 9
 DB 63 ORIKASKSF 71

RESULT 13
 US-08-557-146-2
 ; Sequence 2, Application US/08557146
 ; Patent No. 5834290
 ; GENERAL INFORMATION:
 ; APPLICANT: Egelrud, Torbjorn
 ; APPLICANT: Hanson, Lennart
 ; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
 ; TITLE OF INVENTION: Enzyme (SCCE)
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: White & Case, Patent Department
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036-2787
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/557,146
 ; FILING DATE: 14-DEC-1995
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sterner, Richard J.

REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-181
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-557-145-2

Query Match 100.0%; Score 42; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9
DB 91 QRIKASKSF 99

RESULT 14

US-08-824-874-3
Sequence 3, Application US/08824874
Patent No. 5962300
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
TITLE OF INVENTION: NOVEL KALLIKREIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/824,874
FILING DATE: Filed Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0252 US
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 532504
US-08-824-874-3

Query Match 100.0%; Score 42; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9

Db 91 QRIKASKSF 99

RESULT 15

US-09-154-344-2
Sequence 2, Application US/09154344
Patent No. 5981256
GENERAL INFORMATION:
APPLICANT: Egelrud, Torbjorn
APPLICANT: Hansson, Lennart
TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
TITLE OF INVENTION: Enzyme (SCCE)
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: White & Case, Patent Department
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/154,344
FILING DATE: 16-SEP-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/557,146
FILING DATE: 14-DEC-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sterner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-181
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-154-344-2

Query Match 100.0%; Score 42; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9
DB 91 QRIKASKSF 99

Search completed: March 1, 2004, 17:38:27
Job time : 12.8889 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:16:55 ; Search time 45.5556 Seconds
(without alignments)
55,820 Million cell updates/sec

Title: US-09-905-083-99
Perfect score: 42
Sequence: 1 QRIKASKSF 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues
Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A_Geneseq_29Jan04:*
- 1: Geneseqp1980s:*
 - 2: Geneseqp1990s:*
 - 3: Geneseqp2000s:*
 - 4: Geneseqp2001s:*
 - 5: Geneseqp2002s:*
 - 6: Geneseqp2003as:*
 - 7: Geneseqp2003bs:*
 - 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	9	4	AAE08334 Human str
2	42	100.0	9	4	AAE08334 Human str
3	42	100.0	97	6	ADA05740 Human NOV
4	42	100.0	181	6	ADA05738 Human NOV
5	42	100.0	198	6	ADA05736 Human NOV
6	42	100.0	224	6	ADA05744 Human NOV
7	42	100.0	225	4	AAE05742 Human str
8	42	100.0	247	6	ADA05742 Human NOV
9	42	100.0	250	6	ADA05732 Human NOV
10	42	100.0	252	6	ADA05734 Human NOV
11	42	100.0	253	2	AAE07888 Human str
12	42	100.0	253	2	AAW05383 Human str
13	42	100.0	253	5	ABB84421 Human SCC
14	42	100.0	253	5	ABB84406 Human SCC
15	42	100.0	253	5	AAU82740 Amino aci
16	42	100.0	253	6	ABU07440 Protein d
17	42	100.0	253	6	ABU07471 Protein d
18	42	100.0	253	6	ABR58471 Human str
19	42	100.0	253	7	ADB50484 Ovarian c
20	42	100.0	257	3	AAE21326 Human HSC
21	39	92.9	226	5	ABB84422 Rat SCCE
22	37	88.1	136	4	ABG23378 Novel hum
23	36	85.7	9	4	AAE08302 Human str
24	36	85.7	11	5	ABG68630 Stratum c
25	36	85.7	249	5	ABB84423 Murine SC

ALIGNMENTS

RESULT 1
AAE08334

ID AAE08334 standard; peptide; 9 AA.

XX AC AAE08334;

XX DT 01-NOV-2001 (first entry)

XX DE Human stratum corneum chymotrypsin enzyme peptide #99 (residues 91-99).
XX KW Stratum corneum chymotrypsin enzyme; SCCE; cytostatic; vaccine; tumour;
XX KW cancer; ovarian; breast; lung; colon; prostate; carcinoma; human;
XX KW anticense therapy; malignant hyperplasia.

XX OS Homo sapiens.

XX FN WO200159158-A1.

XX PD 16-AUG-2001.

XX PF 07-FEB-2001; 2001WO-US003977.

XX PR 11-FEB-2000; 2000US-00502600.

XX PA (UYAR-) UNIV ARKANSAS.

XX PI O'brien TJ;

XX DR WPI; 2001-514676/56.

XX PT Diagnosing cancer comprises detecting stratum corneum chymotrypsin enzyme.

XX PS Disclosure; Page 124; 127pp; English.

XX CC The invention relates to diagnosing cancer especially ovarian cancer, by screening for stratum corneum chymotrypsin enzyme (SCCE). Proteases are considered to be an integral part of tumour growth and metastasis, and therefore, markers indicative of their presence or absence are useful for the diagnosis of cancer. The method is useful for diagnosing cancer, carcinoma and malignant hyperplasia. The SCCE oligonucleotide may be used to treat a cancer selected from ovarian, breast, lung, colon, prostate and other cancers in which SCCE is overexpressed. The present sequence is human SCCE peptide

SQ Sequence 9 AA;

Query Match 100.0%; Score 42; DB 4; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.4e+06; Mismatches 0; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9
 |||||
 Db 1 QRIKASKSF 9

RESULT 2
 AA08304
 ID AA08304 standard; peptide; 9 AA.
 XX
 AC AA08304;
 XX
 DT 01-NOV-2001 (first entry)
 XX
 DE Human stratum corneum chymotrypsin enzyme peptide #69 (residues 91-99).
 XX
 KW Stratum corneum chymotrypsin enzyme; SCCE; cytostatic; vaccine; tumour;
 KW cancer; ovarian; breast; lung; colon; prostate; carcinoma; human;
 KW antisense therapy; malignant hyperplasia.
 XX
 OS Homo sapiens.
 XX
 PN WO200159158-A1.
 XX
 PD 16-AUG-2001.
 XX
 PF 07-FEB-2001; 2001WO-US003977.
 XX
 PR 11-FEB-2000; 2000US-00502600.
 XX
 PA (UYAR-) UNIV ARKANSAS.
 XX
 PI O'brien TJ;
 XX
 DR WPI; 2001-514676/56.
 XX
 PT Diagnosing cancer comprises detecting stratum corneum chymotrypsin
 PT enzyme.
 PS Claim 25; Page 117; 127pp; English.
 XX
 CC The invention relates to diagnosing cancer especially ovarian cancer, by
 CC screening for stratum corneum chymotrypsin enzyme (SCCE). Proteases are
 CC considered to be an integral part of tumour growth and metastasis, and
 CC therefore, markers indicative of their presence or absence are useful for
 CC the diagnosis of cancer. The method is useful for diagnosing cancer,
 CC carcinoma and malignant hyperplasia. The SCCE oligonucleotide may be used
 CC to treat a cancer selected from ovarian, breast, lung, colon, prostate
 CC and other cancers in which SCCE is overexpressed. The present sequence is
 CC human SCCE peptide
 XX
 SQ Sequence 9 AA;
 Query Match 100.0%; Score 42; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9
 |||||
 Db 1 QRIKASKSF 9

RESULT 3
 ADA05740
 ID ADA05740 standard; protein; 97 AA.
 XX
 AC ADA05740;
 XX
 DT 06-NOV-2003 (first entry)
 XX
 DE Human NOV18e protein SEQ ID NO:100.

XX human; NOX; antidiabetic; anorectic; antibacterial; virucide;
 KW immunomodulator; cytostatic; nootropic; neuroprotective;
 KW antiparkinsonian; antilipemic; gene therapy; human disease;
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; dyslipidaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO2003029424-A2.
 XX
 PD 10-APR-2003.
 XX
 PF 02-OCT-2002; 2002WO-US031373.
 XX
 PR 02-OCT-2001; 2001US-0326483P.
 PR 05-OCT-2001; 2001US-0327435P.
 PR 05-OCT-2001; 2001US-0327449P.
 PR 09-OCT-2001; 2001US-0327917P.
 PR 09-OCT-2001; 2001US-0328029P.
 PR 09-OCT-2001; 2001US-0328044P.
 PR 09-OCT-2001; 2001US-0328056P.
 PR 12-OCT-2001; 2001US-0328849P.
 PR 15-OCT-2001; 2001US-0329414P.
 PR 17-OCT-2001; 2001US-0330142P.
 PR 18-OCT-2001; 2001US-0330309P.
 PR 22-OCT-2001; 2001US-0341058P.
 PR 24-OCT-2001; 2001US-0339266P.
 PR 24-OCT-2001; 2001US-0343629P.
 PR 29-OCT-2001; 2001US-0349575P.
 PR 01-NOV-2001; 2001US-0346357P.
 PR 17-APR-2002; 2002US-0373260P.
 PR 19-APR-2002; 2002US-0373815P.
 PR 19-APR-2002; 2002US-0373817P.
 PR 19-APR-2002; 2002US-0373826P.
 PR 19-APR-2002; 2002US-0373884P.
 PR 22-APR-2002; 2002US-0374977P.
 PR 16-MAY-2002; 2002US-0381037P.
 PR 16-MAY-2002; 2002US-0381038P.
 PR 16-MAY-2002; 2002US-0381042P.
 PR 17-MAY-2002; 2002US-0381642P.
 PR 28-MAY-2002; 2002US-0383656P.
 PR 29-MAY-2002; 2002US-0383831P.
 PR 23-JUN-2002; 2002US-0391335P.
 PR 01-OCT-2002; 2002US-00262511.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
 PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
 PI Ort T, Gorman L, Zerhusen BD, Anderson DM, Zhong M, Catterton E;
 PI Ji W, Miller CE, Rastell L, Stone DJ, Pena CEA, Shenoy SG;
 PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;
 PI Eissen AJ, Gangolli EA, Rieger DK, Spaderna SK;
 XX
 WPI; 2003-381626/36.
 N-PSDB; ADA05739.
 XX
 DR New NOX polypeptides and nucleic acids, useful for diagnosing,
 PT preventing or treating NOX-associated disorders, e.g. diabetes, obesity,
 PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.
 XX
 PS Claim 1; Page 171; 586pp; English.
 XX
 CC The present invention describes NOX proteins, where X can be 1 to 55
 CC (e.g. NOX1). Also described: (1) a composition comprising a polypeptide
 CC described above and a carrier; (2) a kit comprising, in one or more
 CC containers, the composition described above; (3) an isolated nucleic acid
 CC molecule which encodes a NOX protein of the invention; (4) a vector
 CC comprising the nucleic acid molecule described above; (5) a cell
 CC comprising the above vector; (6) an antibody that immunospecifically

CC binds to the polypeptide described above; (7) methods for determining the
 CC presence or amount of the above polypeptide or nucleic acid molecule in a
 CC sample; (8) methods for determining the presence of or predisposition to
 CC a disease associated with altered levels of expression of the above
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
 CC method of identifying an agent that binds to the polypeptide described
 CC above; (10) a method for identifying a potential therapeutic agent for
 CC use in treating a pathology that is related to an aberrant expression or
 CC aberrant physiological interactions of the polypeptide; (11) a method of
 CC screening for a modulator of activity or of latency or predisposition to
 CC a pathology associated with the polypeptide; (12) a method for modulating
 CC the activity of the polypeptide described above; (13) methods of treating
 CC or preventing a pathology associated with the above polypeptide in a
 CC mammal; and (14) a method for producing the above polypeptide. NOVX
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,
 CC immunomodulator, cytostatic, norectic, neuroprotective, antiparkinsonian
 CC and antilipaeamic activities, and can be used in gene therapy. The
 CC polypeptide is useful in manufacturing a medicament for treating a
 CC syndrome associated with a human disease. The polypeptide or the nucleic
 CC acid molecule may be used to diagnose, treat or prevent metabolic
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
 CC disease, immune disorders, haematopoietic disorders and various
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation
 CC probes, in chromosome mapping, tissue typing, preventive medicine and
 CC pharmacogenomics. The present sequence represents a human NOVX from the
 CC present invention.

XX Sequence 97 AA;

Query Match 100.0%; Score 42; DB 6; Length 97;
 Best Local Similarity 100.0%; Pred. No. 0.19;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ORIKASKSF 9
 Db 72 ORIKASKSF 80

RESULT 4

ADA05738

ID ADA05738 standard; protein; 181 AA.

XX ADA05738;

XX 06-NOV-2003 (first entry)

DE Human NOV18d protein SEQ ID NO:98.

XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
 XX immunomodulator; cytostatic; norectic; neuroprotective;
 XX antiparkinsonian; antilipaeamic; gene therapy; human disease;
 XX metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
 XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 XX immune disorder; haematopoietic disorder; dyslipidaemia.

OS Homo sapiens.

XX WO2003029424-A2.

XX 10-APR-2003.

XX 02-OCT-2002; 2002WO-US031373.

XX 02-OCT-2001; 2001US-0326483P.

XX 05-OCT-2001; 2001US-0327435P.

XX 09-OCT-2001; 2001US-0327917P.

XX 09-OCT-2001; 2001US-0328029P.

XX 09-OCT-2001; 2001US-0328044P.

XX 09-OCT-2001; 2001US-0328056P.

XX 12-OCT-2001; 2001US-0328499P.

XX 15-OCT-2001; 2001US-0329414P.

PR 17-OCT-2001; 2001US-0330142P.
 PR 18-OCT-2001; 2001US-0330309P.
 PR 22-OCT-2001; 2001US-0341058P.
 PR 24-OCT-2001; 2001US-0339266P.
 PR 24-OCT-2001; 2001US-0343629P.
 PR 29-OCT-2001; 2001US-0349575P.
 PR 01-NOV-2001; 2001US-0346357P.
 PR 17-APR-2002; 2002US-0373260P.
 PR 19-APR-2002; 2002US-0373815P.
 PR 19-APR-2002; 2002US-0373817P.
 PR 19-APR-2002; 2002US-0373826P.
 PR 19-APR-2002; 2002US-0373884P.
 PR 22-APR-2002; 2002US-0374977P.
 PR 16-MAY-2002; 2002US-0381037P.
 PR 16-MAY-2002; 2002US-0381038P.
 PR 16-MAY-2002; 2002US-0381042P.
 PR 17-MAY-2002; 2002US-0381842P.
 PR 28-MAY-2002; 2002US-0383856P.
 PR 29-MAY-2002; 2002US-0383831P.
 PR 25-JUN-2002; 2002US-0391335P.
 PR 01-OCT-2002; 2002US-00262511.

(CURA-) CURAGEN CORP.

XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
 XX Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
 XX Ort T, Gorman L, Zerkusen BD, Anderson DW, Zhong M, Catterton E;
 XX Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;
 XX Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;
 XX Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;

XX WPI; 2003-381626/36.

XX N-PSDB; ADA05737.

XX New NOVX polypeptides and nucleic acids, useful for diagnosing,

XX preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
 XX cancer or dyslipidemia, and in chromosome mapping, tissue typing or
 XX pharmacogenomics.

XX Claim 1; Page 171; 586pp; English.

XX The present invention describes NOVX proteins, where X can be 1 to 55
 XX (e.g. NOV1). Also described: (1) a composition comprising a polypeptide
 XX described above and a carrier; (2) a kit comprising, in one or more
 XX containers, the composition described above; (3) an isolated nucleic acid
 XX molecule which encodes a NOVX protein of the invention; (4) a vector
 XX comprising the nucleic acid molecule described above; (5) a cell
 XX comprising the above vector; (6) an antibody that immunospecifically
 XX binds to the polypeptide described above; (7) methods for determining the
 XX presence or amount of the above polypeptide or nucleic acid molecule in a
 XX sample; (8) methods for determining the presence of or predisposition to
 XX a disease associated with altered levels of expression of the above
 XX polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
 XX method of identifying an agent that binds to the polypeptide described
 XX above; (10) a method for identifying a potential therapeutic agent for
 XX use in treating a pathology that is related to an aberrant expression or
 XX aberrant physiological interactions of the polypeptide; (11) a method of
 XX screening for a modulator of activity or of latency or predisposition to
 XX a pathology associated with the polypeptide; (12) a method for modulating
 XX the activity of the polypeptide described above; (13) methods of treating
 XX or preventing a pathology associated with the above polypeptide in a
 XX mammal; and (14) a method for producing the above polypeptide. NOVX
 XX sequences have antidiabetic, anorectic, antibacterial, virucide,
 XX immunomodulator, cytostatic, norectic, neuroprotective, antiparkinsonian
 XX and antilipaeamic activities, and can be used in gene therapy. The
 XX polypeptide is useful in manufacturing a medicament for treating a
 XX syndrome associated with a human disease. The polypeptide or the nucleic
 XX acid molecule may be used to diagnose, treat or prevent metabolic
 XX disorders such as diabetes or obesity, infections, cachexia, cancer,
 XX neurodegenerative disorders such as Alzheimer's disease or Parkinson's
 XX disease, immune disorders, haematopoietic disorders and various
 XX dyslipidaemias. The nucleic acids can also be used as hybridisation
 XX probes, in chromosome mapping, tissue typing, preventive medicine and

CC pharmacogenomics. The present sequence represents a human NOVX from the
 CC present invention.

XX Sequence 181 AA;

Query Match 100.0%; Score 42; DB 6; Length 181;

Best Local Similarity 100.0%; Pred. No. 0.36; 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9

Db 72 QRIKASKSF 80

RESULT 5

ADA05736
 ID ADA05736 standard; protein; 198 AA.

XX AC ADA05736;

XX DT 06-NOV-2003 (first entry)

XX DE Human NOV18c protein SEQ ID NO:96.

XX KW human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
 KW immunomodulator; cytostatic; nontropic; neuroprotective;
 KW antiparkinsonian; antilipemic; gene therapy; human disease;
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; dyslipidaemia.

XX OS Homo sapiens.

XX EN WO2003029424-A2.

XX PD 10-APR-2003.

XX PF 02-OCT-2002; 2002WO-US031373.

XX PR 02-OCT-2001; 2001US-0326483P.

XX PR 05-OCT-2001; 2001US-0327435P.

XX PR 09-OCT-2001; 2001US-0327917P.

XX PR 09-OCT-2001; 2001US-0328029P.

XX PR 09-OCT-2001; 2001US-0328044P.

XX PR 09-OCT-2001; 2001US-0328056P.

XX PR 12-OCT-2001; 2001US-0328849P.

XX PR 15-OCT-2001; 2001US-0329414P.

XX PR 17-OCT-2001; 2001US-0330142P.

XX PR 18-OCT-2001; 2001US-0330309P.

XX PR 22-OCT-2001; 2001US-0341058P.

XX PR 24-OCT-2001; 2001US-0339268P.

XX PR 24-OCT-2001; 2001US-0343629P.

XX PR 29-OCT-2001; 2001US-0349575P.

XX PR 01-NOV-2001; 2001US-0346357P.

XX PR 17-APR-2002; 2002US-0373260P.

XX PR 19-APR-2002; 2002US-0373815P.

XX PR 19-APR-2002; 2002US-0373817P.

XX PR 19-APR-2002; 2002US-0373826P.

XX PR 19-APR-2002; 2002US-0373884P.

XX PR 22-APR-2002; 2002US-0374977P.

XX PR 16-MAY-2002; 2002US-0381037P.

XX PR 16-MAY-2002; 2002US-0381038P.

XX PR 16-MAY-2002; 2002US-0381042P.

XX PR 17-MAY-2002; 2002US-0381642P.

XX PR 28-MAY-2002; 2002US-0383656P.

XX PR 29-MAY-2002; 2002US-0383834P.

XX PR 25-JUN-2002; 2002US-0391335P.

XX PR 01-OCT-2002; 2002US-00262511.

XX (CURA-) CURAGEN CORP.

XX PA Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;

PI

PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
 PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Caterton E;
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;
 PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;
 PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;
 XX WPI: 2003-381626/36.
 DR N-PSDB; ADA05735.

XX New NOVX polypeptides and nucleic acids, useful for diagnosing,
 PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
 PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.

XX Claim 1; Page 170; 586pp; English.

CC The present invention describes NOVX proteins, where X can be 1 to 55
 CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide
 CC described above and a carrier; (2) a kit comprising, in one or more
 CC containers, the composition described above; (3) an isolated nucleic acid
 CC molecule which encodes a NOVX protein of the invention; (4) a vector
 CC comprising the nucleic acid molecule described above; (5) a cell
 CC comprising the above vector; (6) an antibody that immunospecifically
 CC binds to the polypeptide described above; (7) methods for determining the
 CC presence or amount of the above polypeptide or nucleic acid molecule in a
 CC sample; (8) methods for determining the presence of or predisposition to
 CC a disease associated with altered levels of expression of the above
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
 CC method of identifying an agent that binds to the polypeptide described
 CC above; (10) a method for identifying a potential therapeutic agent for
 CC use in treating a pathology that is related to an aberrant expression or
 CC aberrant physiological interactions of the polypeptide; (11) a method of
 CC screening for a modulator of activity or of latency or predisposition to
 CC a pathology associated with the polypeptide; (12) a method for modulating
 CC the activity of the polypeptide described above; (13) methods of treating
 CC or preventing a pathology associated with the above polypeptide in a
 CC mammal; and (14) a method for producing the above polypeptide. NOVX
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,
 CC immunomodulator, cytostatic, nontropic, neuroprotective, antiparkinsonian
 CC and antilipemic activities, and can be used in gene therapy. The
 CC polypeptide is useful in manufacturing a medicament for treating a
 CC syndrome associated with a human disease. The polypeptide or the nucleic
 CC acid molecule may be used to diagnose, treat or prevent metabolic
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
 CC disease, immune disorders, haematopoietic disorders and various
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation
 CC probes, in chromosome mapping, tissue typing, preventive medicine and
 CC pharmacogenomics. The present sequence represents a human NOVX from the
 CC present invention.

XX Sequence 198 AA;

Query Match 100.0%; Score 42; DB 6; Length 198;

Best Local Similarity 100.0%; Pred. No. 0.39;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9

Db 91 QRIKASKSF 99

RESULT 6

ADA05744

ID ADA05744 standard; protein; 224 AA.

XX AC ADA05744;

XX DT 06-NOV-2003 (first entry)

XX DE Human NOV18g protein SEQ ID NO:104.

XX human; NOVX; antidiabetic; anorectic; antibacterial; virucide;

immunomodulator; cytostatic; nootropic; neuroprotective;
 KW antiparkinsonian; antilipemic; gene therapy; human disease;
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
 KW neurodegenerative disorder; Alzheimer's disease; parkinson's disease;
 KW immune disorder; haematopoietic disorder; dyslipidaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO2003029424-A2.
 XX
 PD 10-APR-2003.
 XX
 PF 02-OCT-2002; 2002WO-US031373.
 XX
 PR 02-OCT-2001; 2001US-0326483P.
 PR 05-OCT-2001; 2001US-0327435P.
 PR 05-OCT-2001; 2001US-0327449P.
 PR 09-OCT-2001; 2001US-0327917P.
 PR 09-OCT-2001; 2001US-0328029P.
 PR 09-OCT-2001; 2001US-0328044P.
 PR 09-OCT-2001; 2001US-0328056P.
 PR 12-OCT-2001; 2001US-0328849P.
 PR 12-OCT-2001; 2001US-0329414P.
 PR 17-OCT-2001; 2001US-0330142P.
 PR 18-OCT-2001; 2001US-0330309P.
 PR 22-OCT-2001; 2001US-0341058P.
 PR 24-OCT-2001; 2001US-0339266P.
 PR 24-OCT-2001; 2001US-0343629P.
 PR 29-OCT-2001; 2001US-0349575P.
 PR 01-NOV-2001; 2001US-0346357P.
 PR 17-APR-2002; 2002US-0373260P.
 PR 19-APR-2002; 2002US-0373815P.
 PR 19-APR-2002; 2002US-0373817P.
 PR 19-APR-2002; 2002US-0373826P.
 PR 19-APR-2002; 2002US-0373884P.
 PR 22-APR-2002; 2002US-0374377P.
 PR 16-MAY-2002; 2002US-0381037P.
 PR 16-MAY-2002; 2002US-0381038P.
 PR 16-MAY-2002; 2002US-0381042P.
 PR 17-MAY-2002; 2002US-0381642P.
 PR 28-MAY-2002; 2002US-0383656P.
 PR 28-MAY-2002; 2002US-0383831P.
 PR 25-JUN-2002; 2002US-0391335P.
 PR 01-OCT-2002; 2002US-00262511.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Smithson G, Millet J, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
 PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malvankar UM;
 PI Ort T, Gorman L, Zerkusen BD, Anderson DW, Zhong M, Catterton E;
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;
 PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;
 PI Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;
 XX
 DR WPI: 2003-381626/36.
 DR N-PSDB; ADA05743.
 XX
 XX New NOVX polypeptides and nucleic acids, useful for diagnosing,
 PT preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
 PT cancer or dyslipidaemia, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.
 XX
 XX Claim 1; Page 172; 586pp; English.
 PS
 XX The present invention describes NOVX proteins, where X can be 1 to 55
 CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide
 CC described above and a carrier; (2) a kit comprising, in one or more
 CC containers, the composition described above; (3) an isolated nucleic acid
 CC molecule which encodes a NOVX protein of the invention; (4) a vector
 CC comprising the nucleic acid molecule described above; (5) a cell
 CC comprising the above vector; (6) an antibody that immunospecifically
 CC binds to the polypeptide described above; (7) methods for determining the
 CC presence or amount of the above polypeptide or nucleic acid molecule in a

CC sample; (8) methods for determining the presence of or predisposition to
 CC a disease associated with altered levels of expression of the above
 CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
 CC method of identifying an agent that binds to the polypeptide described
 CC above; (10) a method for identifying a potential therapeutic agent for
 CC use in treating a pathology that is related to an aberrant expression or
 CC aberrant physiological interactions of the polypeptide; (11) a method of
 CC screening for a modulator of activity or of latency or predisposition to
 CC a pathology associated with the polypeptide; (12) a method for modulating
 CC the activity of the polypeptide described above; (13) methods of treating
 CC or preventing a pathology associated with the above polypeptide in a
 CC mammal; and (14) a method for producing the above polypeptide. NOVX
 CC sequences have antidiabetic, anorectic, antibacterial, virucide,
 CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian
 CC and antilipemic activities, and can be used in gene therapy. The
 CC polypeptide is useful in manufacturing a medicament for treating a
 CC syndrome associated with a human disease. The polypeptide or the nucleic
 CC acid molecule may be used to diagnose, treat or prevent metabolic
 CC disorders such as diabetes or obesity, infections, cachexia, cancer,
 CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
 CC disease, immune disorders, haematopoietic disorders and various
 CC dyslipidaemias. The nucleic acids can also be used as hybridisation
 CC probes, in chromosome mapping, tissue typing, preventive medicine and
 CC pharmacogenomics. The present sequence represents a human NOVX from the
 CC present invention.
 XX
 XX Sequence 224 AA;
 SQ
 Query Match 100.0%; Score 42; DB 6; Length 224;
 Best Local Similarity 100.0%; Pred. No. 0.45; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0;
 QY 1 QRIKASKSF 9
 DQ |||||
 Db 77 QRIKASKSF 85
 RESULT 7
 ID AAB98502 standard; protein; 225 AA.
 XX
 AC AAB98502;
 XX
 DT 03-AUG-2001 (first entry)
 XX
 DE Human Stratum Corneum Chymotryptic Enzyme, SCCE, catalytic domain.
 XX
 KW Human; TADG-15; cytostatic; vaccine; ovarian tumour; cancer;
 KW tumour antigen-derived gene 15; serine protease;
 KW Stratum Corneum Chymotryptic Enzyme; SCCE.
 XX
 OS Homo sapiens.
 XX
 PN WO200129056-A1.
 XX
 PD 26-APR-2001.
 XX
 PF 20-OCT-2000; 2000WO-US029095.
 XX
 PR 20-OCT-1999; 99US-00421213.
 XX
 PA (UYAR-) UNIV ARKANSAS.
 XX
 PI O'brien TJ, Tanimoto H;
 XX
 DR WPI; 2001-381031/40.
 XX
 PT Novel extracellular serine protease, termed tumor antigen-derived gene 15
 PT protein overexpressed in carcinomas and DNA encoding it, for diagnosis,
 PT treatment, prevention of cancer, particularly breast, ovarian cancer.
 XX
 PS Example 10; Fig 1; 130pp; English.
 XX

09-OCT-2001; 2001US-0328029P.
09-OCT-2001; 2001US-0328044P.
09-OCT-2001; 2001US-0328056P.
12-OCT-2001; 2001US-0328849P.
15-OCT-2001; 2001US-0329414P.
17-OCT-2001; 2001US-0330142P.
18-OCT-2001; 2001US-0330309P.
22-OCT-2001; 2001US-0341058P.
24-OCT-2001; 2001US-0339266P.
24-OCT-2001; 2001US-0343629P.
29-OCT-2001; 2001US-0349575P.
01-NOV-2001; 2001US-0346357P.
17-APR-2002; 2002US-0373260P.
19-APR-2002; 2002US-0373815P.
19-APR-2002; 2002US-0373817P.
19-APR-2002; 2002US-0373826P.
19-APR-2002; 2002US-0373844P.
22-APR-2002; 2002US-0374977P.
16-MAY-2002; 2002US-0381037P.
16-MAY-2002; 2002US-0381038P.
16-MAY-2002; 2002US-0381042P.
17-MAY-2002; 2002US-0381642P.
28-MAY-2002; 2002US-0383656P.
29-MAY-2002; 2002US-0383831P.
25-JUN-2002; 2002US-0391335P.
01-OCT-2002; 2002US-00262511.
(CURA-) CURAGEN CORP.
Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
Ort T, Gorman L, Zerkhusen BD, Anderson DW, Zhong M, Catterton E;
Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;
Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;
Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;
WPI; 2003-381626/36.
N-PSDB; ADA05733.
New NOVX polypeptides and nucleic acids, useful for diagnosing,
preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
cancer or dyslipidemia, and in chromosome mapping, tissue typing or
pharmacogenomics.
Claim 1; Page 170; 586pp; English.
The present invention describes NOVX proteins, where X can be 1 to 55
(e.g. NOV1). Also described: (1) a composition comprising a polypeptide
described above and a carrier; (2) a kit comprising, in one or more
containers, the composition described above; (3) an isolated nucleic acid
molecule which encodes a NOVX protein of the invention; (4) a vector
comprising the nucleic acid molecule described above; (5) a cell
comprising the above vector; (6) an antibody that immunospecifically
binds to the polypeptide described above; (7) methods for determining the
presence or amount of the above polypeptide or nucleic acid molecule in a
sample; (8) methods for determining the presence of or predisposition to
a disease associated with altered levels of expression of the above
polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
method of identifying an agent that binds to the polypeptide described
above; (10) a method for identifying a potential therapeutic agent for
use in treating a pathology that is related to an aberrant expression or
aberrant physiological interactions of the polypeptide; (11) a method of
screening for a modulator of activity or of latency or predisposition to
a pathology associated with the polypeptide; (12) a method for modulating
the activity of the polypeptide described above; (13) methods of treating
or preventing a pathology associated with the above polypeptide in a
mammal; and (14) a method for producing the above polypeptide. NOVX
sequences have antidiabetic, anorectic, antibacterial, virucide,
immunomodulator, cytostatic, neurotropic, neuroprotective, antiparkinsonian
and antilipemic activities, and can be used in gene therapy. The
polypeptide is useful in manufacturing a medicament for treating a
syndrome associated with a human disease. The polypeptide or the nucleic
acid molecule may be used to diagnose, treat or prevent metabolic

disorders such as diabetes or obesity, infections, cachexia, cancer,
neurodegenerative disorders such as Alzheimer's disease or Parkinson's
disease, immune disorders, haematopoietic disorders and various
dyslipidaemias. The nucleic acids can also be used as hybridisation
probes, in chromosome mapping, tissue typing, preventive medicine and
pharmacogenomics. The present sequence represents a human NOVX from the
present invention.
Sequence 252 AA;
Query Match 100.0%; Score 42; DB 6; Length 252;
Best Local Similarity 100.0%; Pred. No. 0.51;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 QRIKASKSP 9
Db 90 QRIKASKSF 98
RESULT 11
AAR67888
ID AAR67888 standard; protein; 253 AA.
XX AAR67888;
AC AAR67888;
XX 25-MAR-2003 (revised)
DT 09-AUG-1995 (first entry)
XX XX
DE Human stratum corneum chymotrophic recombinant enzyme (SCCE).
XX Stratum corneum chymotryptic enzyme; skin disorder; acne; psoriasis;
KW callosities; keratosis pilaris; ichthyoses; eczema.
XX Homo sapiens.
XX WO9500651-A1.
XX XX
PD 05-JAN-1995.
XX 20-JUN-1994; 94WO-IB000166.
XX 18-JUN-1993; 93DK-00000725.
XX (SYMB-) SYMBICOM AB.
XX Egelrud T, Hansson L;
XX WPI; 1995-052088/07.
XX N-PSDB; AAQ81203.
XX Nucleotide sequences encoding stratum corneum chymotryptic enzyme - and
PT related vectors, transformed cells and polypeptides, useful for treating
PT skin disorders, e.g. acne or psoriasis, and for identification of
PT specific inhibitors.
XX Disclosure; Page 97; 137pp; English.
XX The enzyme encoded by this sequence is used in pharmaceutical, cosmetic
CC and skin care products, especially to treat and prevent acne, xeroderma,
CC or other hyperkeratotic conditions (e.g. callosities or keratosis
CC pilaris), ichthyoses, psoriasis, eczema, etc. It is produced
CC recombinantly following mammal, insect, plant, or microorganism
CC transformation with plasmid pS507. (Updated on 25-MAR-2003 to correct PN
CC field.)
XX Sequence 253 AA;
Query Match 100.0%; Score 42; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.51;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 QRIKASKSP 9
Db 90 QRIKASKSF 98

Db	91 QRIKASKSF 99	
XX		
OS	Homo sapiens.	
XX		
PN	WO200262135-A2.	
XX		
PD	15-AUG-2002.	
XX		
PF	08-FEB-2002; 2002WO-IB001300.	
XX		
PR	09-FEB-2001; 2001CA-02332655.	
XX		
PR	09-FEB-2001; 2001EK-00000218.	
XX		
PA	(EGBL/) EGBLRUD T.	
XX	(HANS/) HANSSON L.	
XX		
PI	Egelrud T, Hansson L;	
XX		
XX	WPI; 2002-643380/69.	
XX		
DR	Transgenic mammal or its embryo useful as model for human disease, has	
XX	heterologous nucleotide sequence coding for stratum corneum chymotryptic	
PT	enzyme operably linked to promoter that drives its expression in skin.	
PT		
XX	Example 6; Page 37; 74pp; English.	
XX		
CC	This invention describes a novel non-human transgenic mammal or mammalian	
CC	embryo having integrated within its genome, a heterologous nucleotide	
CC	sequence comprising at least a significant part of a nucleotide sequence	
CC	coding for a stratum corneum chymotryptic enzyme (SCCE) or its variant.	
CC	operably linked to a promoter that drives expression of heterologous SCCE	
CC	or its variant in skin. The product of the invention is useful as a model	
CC	for the study of disease with the aim of improving treatment, to relieve	
CC	or ameliorate a pathogenic condition, for development or testing of a	
CC	cosmetic or a pharmaceutical formulation, and for the development of a	
CC	diagnostic method. It can also be used as a model for a skin disease or	
CC	skin cancer. The invention is also useful for screening or identifying a	
CC	compound or composition effective for the prevention or treatment of an	
CC	abnormal or unwanted phenotype, and for screening or identifying a	
CC	compound or composition effective for the prevention or treatment of	
CC	inflammatory skin diseases selected from diseases consisting of epidermal	
CC	hyperkeratosis, acanthosis, epidermal inflammation, dermal inflammation,	
CC	pruritus, atopic dermatitis, eczema, acne and inherited skin diseases	
CC	with epidermal hyperkeratosis. The mammal of the invention is also useful	
CC	as a model for further studies of itch mechanisms and the testing of	
CC	potential compounds and compositions for relieve of various skin diseases	
CC	where itch is a component. This sequence represents the N-terminal	
CC	fragment of the human stratum corneum chymotryptic enzyme, SCCE	
CC	synonymous with human kallikrein 7 (KLK7), used in the development of the	
CC	transgenic mammals described in the invention	
XX		
SQ	Sequence 253 AA;	
	Query Match 100.0%; Score 42; DB 5; Length 253;	
	Best Local Similarity 100.0%; Pred. No. 0.51;	
	Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 QRIKASKSF 9	
Db	91 QRIKASKSF 99	
RESULT 13		
ABB84421		
ID	ABB84421 standard; peptide; 253 AA.	
XX		
AC	ABB84421;	
XX		
DT	08-NOV-2002 (first entry)	
XX		
DE	Human SCCE protein N-terminal fragment SEQ ID 48.	
XX		
XX	SCCE; human; stratum corneum chymotryptic enzyme; kallikrein 7;	
KW	serine protease; transgenic mammal; skin; skin disease; skin cancer;	
KW	hyperkeratosis; acanthosis; epidermal inflammation; dermal inflammation;	
KW	pruritus; atopic dermatitis; eczema; acne; itch; KLK7; enzyme.	

XX		
OS	Homo sapiens.	
XX		
PN	WO200262135-A2.	
XX		
PD	15-AUG-2002.	
XX		
PF	08-FEB-2002; 2002WO-IB001300.	
XX		
PR	09-FEB-2001; 2001CA-02332655.	
XX		
PR	09-FEB-2001; 2001EK-00000218.	
XX		
PA	(EGBL/) EGBLRUD T.	
XX	(HANS/) HANSSON L.	
XX		
PI	Egelrud T, Hansson L;	
XX		
XX	WPI; 2002-643380/69.	
XX		
DR	Transgenic mammal or its embryo useful as model for human disease, has	
XX	heterologous nucleotide sequence coding for stratum corneum chymotryptic	
PT	enzyme operably linked to promoter that drives its expression in skin.	
PT		
XX	Example 6; Page 37; 74pp; English.	
XX		
CC	This invention describes a novel non-human transgenic mammal or mammalian	
CC	embryo having integrated within its genome, a heterologous nucleotide	
CC	sequence comprising at least a significant part of a nucleotide sequence	
CC	coding for a stratum corneum chymotryptic enzyme (SCCE) or its variant.	
CC	operably linked to a promoter that drives expression of heterologous SCCE	
CC	or its variant in skin. The product of the invention is useful as a model	
CC	for the study of disease with the aim of improving treatment, to relieve	
CC	or ameliorate a pathogenic condition, for development or testing of a	
CC	cosmetic or a pharmaceutical formulation, and for the development of a	
CC	diagnostic method. It can also be used as a model for a skin disease or	
CC	skin cancer. The invention is also useful for screening or identifying a	
CC	compound or composition effective for the prevention or treatment of an	
CC	abnormal or unwanted phenotype, and for screening or identifying a	
CC	compound or composition effective for the prevention or treatment of	
CC	inflammatory skin diseases selected from diseases consisting of epidermal	
CC	hyperkeratosis, acanthosis, epidermal inflammation, dermal inflammation,	
CC	pruritus, atopic dermatitis, eczema, acne and inherited skin diseases	
CC	with epidermal hyperkeratosis. The mammal of the invention is also useful	
CC	as a model for further studies of itch mechanisms and the testing of	
CC	potential compounds and compositions for relieve of various skin diseases	
CC	where itch is a component. This sequence represents the N-terminal	
CC	fragment of the human stratum corneum chymotryptic enzyme, SCCE	
CC	synonymous with human kallikrein 7 (KLK7), used in the development of the	
CC	transgenic mammals described in the invention	
XX		
SQ	Sequence 253 AA;	
	Query Match 100.0%; Score 42; DB 2; Length 253;	
	Best Local Similarity 100.0%; Pred. No. 0.51;	
	Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 QRIKASKSF 9	
Db	91 QRIKASKSF 99	
RESULT 14		
ABB84406		
ID	ABB84406 standard; protein; 253 AA.	
XX		
AC	ABB84406;	
XX		
DT	08-NOV-2002 (first entry)	
XX		
DE	Human SCCE protein.	
XX		
XX	SCCE; human; stratum corneum chymotryptic enzyme; kallikrein 7;	
KW	serine protease; transgenic mammal; skin; skin disease; skin cancer;	
KW	hyperkeratosis; acanthosis; epidermal inflammation; dermal inflammation;	
KW	pruritus; atopic dermatitis; eczema; acne; itch; KLK7; enzyme.	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:35:01 ; Search time 24.1111 Seconds

(without alignments)
78.818 Million cell updates/sec

Title: US-09-905-083-99

Perfect score: 42

Sequence: 1 QRIKASKSF 9

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	100.0	9	9	US-09-918-243-99
2	42	100.0	9	9	Sequence 99, Appl
3	42	100.0	9	9	Sequence 130, Appl
4	42	100.0	9	9	Sequence 99, Appl
5	42	100.0	144	9	US-09-905-083-130
6	42	100.0	144	9	Sequence 4, Appl
7	42	100.0	253	14	US-10-461-787-4
8	42	100.0	253	9	Sequence 98, Appl
9	42	100.0	253	14	Sequence 3, Appl
10	42	100.0	253	14	Sequence 90, Appl
11	42	100.0	253	15	US-10-264-762-3
12	36	85.7	9	9	Sequence 498, Appl
13	36	85.7	9	9	Sequence 48, Appl
14	33	78.6	427	10	US-09-918-243-97
15	33	78.6	427	14	Sequence 97, Appl

16	33	78.6	427	14	US-10-005-956-574	Sequence 574, App
17	33	78.6	427	14	US-10-005-956-576	Sequence 576, App
18	33	78.6	427	14	US-10-005-956-578	Sequence 578, App
19	33	78.6	427	14	US-10-005-956-853	Sequence 853, App
20	33	78.6	427	14	US-10-005-956-855	Sequence 855, App
21	31	73.8	35	10	US-09-809-391-619	Sequence 619, App
22	31	73.8	35	10	US-09-882-171-619	Sequence 619, App
23	31	73.8	276	15	US-10-108-260A-3762	Sequence 3762, App
24	31	73.8	298	15	US-10-168-659-20	Sequence 20, Appl
25	31	73.8	319	15	US-10-389-493-2099	Sequence 2099, App
26	31	73.8	1201	10	US-09-863-776-61	Sequence 61, Appl
27	31	73.8	2861	15	US-10-374-979-108	Sequence 108, App
28	31	73.8	2861	15	US-10-331-496A-89	Sequence 89, Appl
29	31	73.8	3038	10	US-09-863-776-62	Sequence 62, Appl
30	30	71.4	627	15	US-10-369-493-18492	Sequence 18492, A
31	29	69.0	46	10	US-09-764-891-2753	Sequence 2753, App
32	29	69.0	127	14	US-10-097-111-333	Sequence 333, App
33	29	69.0	187	13	US-10-062-254-168	Sequence 168, App
34	29	69.0	200	13	US-10-062-254-170	Sequence 170, App
35	29	69.0	205	13	US-10-062-254-130	Sequence 130, App
36	29	69.0	206	13	US-10-062-254-128	Sequence 128, App
37	29	69.0	438	10	US-09-769-736-137	Sequence 137, App
38	28	66.7	61	9	US-09-912-020-357	Sequence 357, App
39	28	66.7	84	9	US-09-864-761-42601	Sequence 42601, A
40	28	66.7	95	9	US-09-867-550-1326	Sequence 1326, App
41	28	66.7	240	11	US-09-833-245-284	Sequence 284, App
42	28	66.7	265	15	US-10-374-780A-1929	Sequence 1929, App
43	28	66.7	306	15	US-10-369-493-1389	Sequence 1389, App
44	28	66.7	374	15	US-10-369-493-11176	Sequence 11176, A
45	28	66.7	416	15	US-10-369-493-1521	Sequence 1521, App

ALIGNMENTS

RESULT 1

US-09-918-243-99
; Sequence 99, Application US/09918243
; Patent No. US20020142317A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/C/D/CIP
; CURRENT APPLICATION NUMBER: US/09/918,243
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 99
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CHAIN
; OTHER INFORMATION: Residues 91-99 of the SCCE protein
US-09-918-243-99

Query Match	100.0%	Score 42;	DB 9;	Length 9;
Best Local Similarity	100.0%	Pred. No. 7.1e+05;		
Matches	9;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;
QY	1	QRIKASKSF	9	
Db	1	QRIKASKSF	9	
RESULT 2				
US-09-918-243-130				
; Sequence 130, Application US/09918243				
; Patent No. US20020142317A1				
; GENERAL INFORMATION:				

NAME/KEY: DOMAIN
OTHER INFORMATION: Amino acid sequence of stratum corneum chymotryptic
OTHER INFORMATION: enzyme (scce) catalytic domain
US-10-461-787-4

Query Match 100.0%; Score 42; DB 14; Length 144;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRIKASKSF 9
Db 28 QRIKASKSF 36

RESULT 7
US-09-888-615-98
Sequence 98, Application US/09888615
Patent No. US20020064858A1
GENERAL INFORMATION:
APPLICANT: FLOWMAN, GREGORY
APPLICANT: WHYTE, DAVID
APPLICANT: CAENEPEL, SEAN
APPLICANT: CHARYDCZAK, GLEN
APPLICANT: MANNING, GERARD
APPLICANT: SUDARSANAM, SUCHA
TITLE OF INVENTION: NOVEL PROTEASES
FILE REFERENCE: 038602/1214
CURRENT APPLICATION NUMBER: US/09/888,615
PRIOR FILING DATE: 2001-06-26
PRIOR APPLICATION NUMBER: 60/214,047
PRIOR FILING DATE: 2000-06-26
NUMBER OF SEQ ID NOS: 150
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 98
LENGTH: 253
TYPE: PRT
ORGANISM: Homo sapiens
US-09-888-615-98

Query Match 100.0%; Score 42; DB 9; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRIKASKSF 9
Db 91 QRIKASKSF 99

RESULT 8
US-09-764-762-3
Sequence 3, Application US/09764762
Patent No. US20020068341A1
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
Lal, Preeti
TITLE OF INVENTION: NOVEL KALLIKREIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/764,762
FILING DATE: 16-Jan-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/210,084
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0252 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 532504
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-764-762-3

Query Match 100.0%; Score 42; DB 9; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRIKASKSF 9
Db 91 QRIKASKSF 99

RESULT 9
US-10-264-283-90
Sequence 90, Application US/10264283
Publication No. US2003014494A1
GENERAL INFORMATION:
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
FILE REFERENCE: 210121.590
CURRENT APPLICATION NUMBER: US/10/264,283
CURRENT FILING DATE: 2002-10-02
NUMBER OF SEQ ID NOS: 111
SOFTWARE: Corixa Invention Disclosure Database
SEQ ID NO 90
LENGTH: 253
TYPE: PRT
ORGANISM: Homo sapiens
US-10-264-283-90

Query Match 100.0%; Score 42; DB 14; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRIKASKSF 9
Db 91 QRIKASKSF 99

RESULT 10
US-10-235-027-498
Sequence 498, Application US/10295027
Publication No. US20030232350A1
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezi, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard

DB 2 QRIKASKS 9

RESULT 14

US-09-993-180-4
; Sequence 4, Application US/09993180
; Publication No. US20030054445A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN SERPIN SECRETED FROM LYMPHOCYTES
; FILE REFERENCE: D0051.NP
; CURRENT APPLICATION NUMBER: US/09/993,180
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/248,434
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 60/257,610
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 60/282,745
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-993-180-4

Query Match 78.6%; Score 33; DB 10; Length 427;
Best Local Similarity 66.7%; Pred. No. 96;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9

DB 355 QKLEASKSF 363

RESULT 15

US-10-005-956-572
; Sequence 572, Application US/10005956
; Publication No. US20030113726A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
; FILE REFERENCE: D0053NP
; CURRENT APPLICATION NUMBER: US/10/005,956
; CURRENT FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: 60/251,015
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 60/263,678
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/273,037
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 1579
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 572
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-005-956-572

Query Match 78.6%; Score 33; DB 14; Length 427;
Best Local Similarity 66.7%; Pred. No. 96;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9

DB 355 QKLEASKSF 363

Search completed: March 1, 2004, 18:08:54
Job time : 25.1111 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:21:01 ; Search time 10.2222 Seconds
(without alignments)
84.690 Million cell updates/sec

Title: US-09-905-083-99

Perfect score: 42

Sequence: 1 QRIKASKSF 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:**

1: piri:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	100.0	253	2 A53968	serine proteinase
2	34	81.0	373	2 D89004	protein T24A6.11 [
3	33	78.6	427	2 A49518	kallistatin precu
4	32	76.2	895	2 D72571	probable DNA-direc
5	31	73.8	248	2 A40625	tetrachloro-p-hydr
6	31	73.8	298	2 S32368	beta-SNAP protein
7	31	73.8	302	2 T50147	max16-like protein
8	31	73.8	362	2 F81409	probable phosphate
9	30	71.4	142	2 C39610	BET1 protein - yea
10	30	71.4	321	2 C71463	probable p-loop Ar
11	30	71.4	331	2 H97051	probable dehydroge
12	30	71.4	427	2 T48159	hypothetical prote
13	30	71.4	517	2 T40129	hypothetical prote
14	30	71.4	627	2 C86850	neutral endopeptid
15	30	71.4	627	2 A47098	lactococcal endope
16	30	71.4	627	2 F53290	endopeptidase P80
17	30	71.4	875	2 T19678	hypothetical prote
18	30	71.4	985	2 D82776	pyruvate dehydroge
19	30	71.4	1102	2 H84545	probable ubiquitin
20	30	71.4	3660	1 S02041	dystrophin, muscle
21	29	69.0	65	2 AF2269	hypothetical prote
22	29	69.0	141	1 B43222	16k protein - toba
23	29	69.0	172	2 H81251	NADH2 dehydrogenas
24	29	69.0	182	2 C64176	hypothetical prote
25	29	69.0	203	2 C90678	thiogalactoside ac
26	29	69.0	203	2 G85528	thiogalactoside ac
27	29	69.0	268	2 E84339	quinolinate phosph
28	29	69.0	349	2 AE3568	WD repeat protein
29	29	69.0	368	2 T23587	hypothetical prote

30	29	69.0	463	2 G97300	aspartyl/asparagin
31	29	69.0	484	2 JCS282	phosphoglucosylat
32	29	69.0	526	1 G71081	probable helicase
33	29	69.0	563	2 B82883	hypothetical prote
34	29	69.0	584	2 T48273	hypothetical prote
35	29	69.0	585	2 I37216	calicin - human (f
36	29	69.0	669	2 AB2028	hypothetical prote
37	29	69.0	691	2 T46476	hypothetical prote
38	29	69.0	717	2 AC1419	DNA topoisomerase
39	29	69.0	738	2 B86294	hypothetical prote
40	29	69.0	743	2 C86168	hypothetical prote
41	29	69.0	759	2 S67164	probable membrane
42	29	69.0	940	2 A40985	projectin - fruit
43	29	69.0	1018	2 T15297	hypothetical prote
44	29	69.0	1742	2 S24600	projectin - fruit
45	29	69.0	6658	2 T13931	projectin - fruit

ALIGNMENTS

RESULT 1

A53968
serine proteinase SCCE precursor - human
N:Alternate names: stratum corneum chymotryptic enzyme
C:Species: Homo sapiens (man)
C:Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 22-Jun-1999
C:Accession: A53968
R:Hansson, L.; Stroenqvist, M.; Baeckman, A.; Wallbrandt, P.; Carlstein, A.; Egelrud, T.
J. Biol. Chem. 269, 19420-19426, 1994
A:Title: Cloning, expression, and characterization of stratum corneum chymotryptic enzym
A:Reference number: A53968; MUID:94308225; PMID:8034709
A:Accession: A53968
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-253 <HAN>
A:Cross-references: GB:I33404; NID:G521214; PIDN:AAC37551.1; PID:G532504
C:Genetics:
A:Gene: GDB:PRSS6; SCCE
A:Cross-references: GDB:377730
A:Map position: 7q35-7q35
C:Superfamily: trypsin; trypsin homology
F:30-245/Domain: trypsin homology <TRY>

Query Match 100.0%; Score 42; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.084;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9

Db 91 QRIKASKSF 99

RESULT 2

D89004
protein T24A6.11 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: D89004
R:anonymous, the C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MUID:9069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: D89004
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-373 <STO>
A:Cross-references: GB:Chr V; PIDN:AAC17798.1; PID:g3168946; GSPDB:GN00023; CESP:T24A6.1
C:Genetics:
A:Gene: T24A6.11

Db 19 KRVKASHSP 27
:|||||

RESULT 7

T50147
mak16-like protein [imported] - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 21-Jul-2000
R:Accession: T50147
R:Hamilton, N.; Churcher, C.M.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, November 1999
A:Reference number: Z25043
A:Accession: T50147
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-302 <HAM>
A:Cross-references: EMBL:AL137798; PIDN:CAB60698.1; GSPDB:GN000666; SPDB:SPAC222.06
A:Experimental source: strain 972h(-); cosmid c222
C:Genetics:
A:Gene: SPDB:SPAC222.06
A:Map position: 1
A:Introns: 21/2
C:Superfamily: MAK16 protein

Query Match 73.8%; Score 31; DB 2; Length 302;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ORIKASKSP 9
:|||||
Db 78 ORIKLSKNY 86

RESULT 8

F81409
probable phosphate transport system permease protein Cj0615 [imported] - Campylobacter jejuni
C:Species: Campylobacter jejuni
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
R:Accession: F81409
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrett, N.; et al.
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypoxanthine auxotrophy
A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: F81409
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-362 <PAR>
A:Cross-references: GB:AL139075; GB:AL111168; NID:G6967817; PIDN:CAB75251.1; PID:G696808
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: pstA; Cj0615

Query Match 73.8%; Score 31; DB 2; Length 362;
Best Local Similarity 77.8%; Pred. No. 40;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ORIKASKSP 9
:|||||
Db 7 KRQKASKSP 15

RESULT 9

C39610
BET1 protein - Yeast (Saccharomyces cerevisiae)
N/Alternate names: protein YIA4c; protein YIL004c; SLY12 protein
C:Species: Saccharomyces cerevisiae
C>Date: 13-Sep-1991 #sequence_revision 13-Sep-1991 #text_change 29-Oct-1999
R:Accession: C39610; S48453; S58677
R:Pascher, C.; Ossig, R.; Gallwitz, D.; Schmitt, H.D.
Mol. Cell. Biol. 11, 872-885, 1991
A:Title: Identification and structure of four yeast genes (SLY) that are able to suppress

A:Reference number: A39610; MUID:91117254; PMID:1990290
A:Accession: C39610
A:Molecule type: DNA
A:Residues: 1-142 <DAS>
A:Cross-references: EMBL:X54237; NID:G4483; PIDN:CAA38143.1; PID:G4484
R:Rowley, N.

submitted to the EMBL Data Library, August 1994

A:Reference number: S48453
A:Accession: S48453
A:Molecule type: DNA
A:Residues: 1-142 <ROW>
A:Cross-references: GB:Z47047; EMBL:Z38113; NID:G603997; PID:G763342; MIPS:YIL004c
R:Voss, H.; Tamames, J.; Teodoru, C.; Valencia, A.; Senses, C.; Wiemann, S.; Schwager, C.
Yeast 11, 61-78, 1995
A:Title: Nucleotide sequence and analysis of the centromeric region of yeast chromosome 1
A:Reference number: S50795; MUID:95282515; PMID:7762303
A:Accession: S58677
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-142 <VOS>
A:Cross-references: EMBL:X79743
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1994
C:Genetics:

A:Gene: SGP:BET1; SLY12
A:Cross-references: SGD:S0001266; MIPS:YIL004c
A:Map position: 9L
A:Introns: 4/2
C:Keywords: transmembrane protein
F:126-142/Domain: transmembrane #status predicted <TMM>

Query Match 71.4%; Score 30; DB 2; Length 142;
Best Local Similarity 87.5%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ORIKASKS 8
:|||||
Db 67 QRIKALKS 74

RESULT 10

C71463
probable p-loop ATPase - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C:Species: Chlamydia trachomatis
C>Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999
C:Accession: C71463
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, S.; et al.
Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis
A:Reference number: A71570; MUID:99000809; PMID:9784136
A:Accession: C71463
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-321 <ARN>
A:Cross-references: GB:AE001357; GB:AE001273; NID:G3329311; PIDN:AAC68437.1; PID:G332931
A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
A:Gene: mesJ

Query Match 71.4%; Score 30; DB 2; Length 321;
Best Local Similarity 66.7%; Pred. No. 60;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ORIKASKSF 9
:|||||
Db 277 ERLVASKSF 285

RESULT 11

H97051
probable dehydrogenase, YUUF B. subtilis ortholog [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C:Accession: H97051

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:17:25 ; Search time 6 seconds
(without alignments)
78.105 Million cell updates/sec

Title: US-09-905-083-99

Perfect score: 42

Sequence: 1 QRIKASKSF 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	100.0	253	1	KLK7 HUMAN
2	34	81.0	107	1	Q9CJU4 pasteurilla
3	33	78.6	314	1	RLA0 THEAC
4	33	78.6	427	1	KAIN HUMAN
5	32	76.2	161	1	RL10 MYCPE
6	31	73.8	247	1	PCPC SPHR
7	31	73.8	298	1	SNAB BOVIN
8	31	73.8	298	1	SNAB HUMAN
9	31	73.8	3038	1	TRIO HUMAN
10	30	71.4	142	1	BET1 YEAST
11	30	71.4	517	1	YB3C SCHPO
12	30	71.4	626	1	PEPO_LACLA
13	30	71.4	626	1	PEPO_LACLC
14	30	71.4	727	1	PKP1 BOVIN
15	30	71.4	728	1	PKP1 MOUSE
16	30	71.4	747	1	PKP1 HUMAN
17	30	71.4	3660	1	DMD CHICK
18	29	69.0	141	1	ORN HAEN
19	29	69.0	182	1	VI6K TRVPL
20	29	69.0	336	1	GPDA HAEDU
21	29	69.0	463	1	SYN_CLOAB
22	29	69.0	484	1	PGPD ACTAC
23	29	69.0	588	1	CALI HUMAN
24	29	69.0	738	1	CNG7 ARATH
25	29	69.0	761	1	YK22 CAEEL
26	29	69.0	853	1	DIA3 HUMAN
27	28	66.7	46	1	DIUH LOEMI
28	28	66.7	61	1	CSRA_ECOLI
29	28	66.7	61	1	CSRA_ERWCA
30	28	66.7	61	1	CSRA_PHOLL
31	28	66.7	69	1	CSRA_SERMA
32	28	66.7	153	1	RS16 BIFLO
33	28	66.7	156	1	GREB_BACSU

RESULT 1
KLK7_HUMAN STANDARD; PRT; 253 AA.
ID KLK7_HUMAN
AC P49862;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Kallikrein 7 precursor (EC 3.4.21.-) (Stratum corneum chymotryptic enzyme) (hSCCE).
DE KLK7 OR PRSS6 OR SCCE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A., AND SEQUENCE OF 23-53.
RP TISSUE=Skin;
RX MEDLINE=94308225; PubMed=8034709;
RA Hansson L., Stroemqvist M., Baeckman A., Wallbrandt P., Carlstein A., Egelrud T.;
RT "Cloning, expression, and characterization of stratum corneum chymotryptic enzyme. A skin-specific human serine proteinase.";
RL J. Biol. Chem. 269:19420-19426(1994).
[2]
RN SEQUENCE FROM N.A.
RP TISSUE=Keratinocytes;
RC Yousef G.M., Scorrilas A., Diamandis E.P.;
RT "Molecular characterization, mapping and tissue expression of the human stratum corneum chymotryptic enzyme gene.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RP MEDLINE=20510030; PubMed=11054574;
RX Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J., Moss P., Paepker B., Wang K.;
RT "Sequencing and expression analysis of the serine protease gene cluster located in chromosome 19q13 region.";
RL Gene 257:119-130(2000).
[4]
RN SEQUENCE FROM N.A.
RP Hansson L., Baeckman A., Ny A., Edlund M., Edholm E., Tornell J., Wallbrandt P., Egelrud T.;
RT "Epidermal overexpression of stratum corneum chymotryptic enzyme in mice; a model for chronic itchy dermatitis.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
[5]
RN CHARACTERIZATION.
RP MEDLINE=95314630; PubMed=7794273;
RX Skvitt A., Stroemqvist M., Egelrud T.;
RT "Primary substrate specificity of recombinant human stratum corneum chymotryptic enzyme.";
RL Biochem. Biophys. Res. Commun. 211:586-589(1995).
CC -!- FUNCTION: May catalyze the degradation of intercellular cohesive structures in the cornified layer of the skin in the continuous shedding of cells from the skin surface. Specific for amino acid residues with aromatic side chains in the P1 position. SCCE

Q811k9 bacillus an
Q817z6 bacillus ce
Q9kdd7 bacillus ha
Q92bl7 listeria mo
Q97jn2 clostridium
P10962 saccharomyc
Q9ut51 schizosacch
P57153 buchnera ap
P00560 saccharomyc
P36661 escherichia
P91343 caenorhabdi
Q9chue lactococcus

ALIGNMENTS

DR EMBL; AL445064; CAC11503.1; --
 DR HAMAP; MF_00280; --; 1.
 DR InterPro; IPR001790; Ribosomal L10.
 DR Pfam; PF00466; Ribosomal L10; 1.
 DR Ribosomal protein; Complete proteome
 KW SEQUENCE 314 AA; 34583 MW; F40D048680B0B6EB CRC64;
 SQ
 Query Match 78.6%; Score 33; DB 1; Length 314;
 Best Local Similarity 87.5%; Pred. No. 5.5;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRIKASKS 8
 DB 17 QRIKASRS 24
 RESULT 4
 KAIN HUMAN
 ID KAIN HUMAN STANDARD; PRT; 427 AA.
 AC F29622; Q96825;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Kallistatin precursor (Kallikrein inhibitor) (Protease inhibitor 4).
 GN SERPINA4 OR P14 OR KST.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94043294; PubMed=8227002;
 RA Chai K.X., Chen L.-M., Chao J., Chao L.;
 RT "Kallistatin: a novel human serine proteinase inhibitor. Molecular
 RT cloning, tissue distribution, and expression in *Escherichia coli*.";
 RL J. Biol. Chem. 268:24498-24505(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95137583; PubMed=7835886;
 RA Chai K.X., Ward D.C., Chao J., Chao L.;
 RT "Molecular cloning, sequence analysis, and chromosomal localization
 RT of the human protease inhibitor 4 (kallistatin) gene (P14).";
 RL Genomics 23:370-378(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Colon;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Teisgold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Boak S.A., McEwan K.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE OF 388-403.
 RC TISSUE=Plasma;
 RX MEDLINE=93100304; PubMed=1334488;
 RA Zhou G.X., Chao L., Chao J.;
 RT "Kallistatin: a novel human tissue kallikrein inhibitor."

RT EMBL; AL445064; CAC11503.1; --
 RN HAMAP; MF_00280; --; 1.
 RP InterPro; IPR001790; Ribosomal L10.
 RX Pfam; PF00466; Ribosomal L10; 1.
 RT Ribosomal protein; Complete proteome
 KW SEQUENCE 314 AA; 34583 MW; F40D048680B0B6EB CRC64;
 SQ
 Query Match 78.6%; Score 33; DB 1; Length 314;
 Best Local Similarity 87.5%; Pred. No. 5.5;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRIKASKS 8
 DB 17 QRIKASRS 24
 RESULT 4
 KAIN HUMAN
 ID KAIN HUMAN STANDARD; PRT; 427 AA.
 AC F29622; Q96825;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Kallistatin precursor (Kallikrein inhibitor) (Protease inhibitor 4).
 GN SERPINA4 OR P14 OR KST.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94043294; PubMed=8227002;
 RA Chai K.X., Chen L.-M., Chao J., Chao L.;
 RT "Kallistatin: a novel human serine proteinase inhibitor. Molecular
 RT cloning, tissue distribution, and expression in *Escherichia coli*.";
 RL J. Biol. Chem. 268:24498-24505(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95137583; PubMed=7835886;
 RA Chai K.X., Ward D.C., Chao J., Chao L.;
 RT "Molecular cloning, sequence analysis, and chromosomal localization
 RT of the human protease inhibitor 4 (kallistatin) gene (P14).";
 RL Genomics 23:370-378(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Colon;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Teisgold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Boak S.A., McEwan K.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE OF 388-403.
 RC TISSUE=Plasma;
 RX MEDLINE=93100304; PubMed=1334488;
 RA Zhou G.X., Chao L., Chao J.;
 RT "Kallistatin: a novel human tissue kallikrein inhibitor."

Purification, characterization, and reactive center sequence.";
 J. Biol. Chem. 267:25873-25880(1992).
 [5]
 CARBOHYDRATE-LINKAGE SITE ASN-157.
 MEDLINE=22660472; PubMed=12754519;
 Zhang H., Li X.-J., Martin D.B., Aebersold R.;
 RT Identification and quantification of N-linked glycoproteins using
 RT hydrazide chemistry, stable isotope labeling and mass spectrometry.";
 Nat. Biotechnol. 21:660-666(2003)
 CC -|- FUNCTION: INHIBITS HUMAN AMIDOLYTIC AND KININOGENASE ACTIVITIES OF
 CC HUMAN TISSUE KALLIKREIN. INHIBITION IS ACHIEVED BY FORMATION OF AN
 CC EQUIMOLAR, HEAT- AND SDS-STABLE COMPLEX BETWEEN THE INHIBITOR AND
 CC THE ENZYME, AND GENERATION OF A SMALL C-TERMINAL FRAGMENT OF THE
 CC INHIBITOR DUE TO CLEAVAGE AT THE REACTIVE SITE BY TISSUE
 CC KALLIKREIN.
 CC -|- SUBUNIT: MONOMER AND SOME HOMODIMERS.
 CC -|- SUBCELLULAR LOCATION: Secreted.
 CC -|- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.
 CC -|- PTM: The N-terminus is blocked.
 CC -|- MISCELLANEOUS: HEPARIN BLOCKS KALLISTATIN'S COMPLEX FORMATION WITH
 CC TISSUE KALLIKREIN AND ABOLISHES ITS INHIBITORY EFFECT ON TISSUE
 CC KALLIKREIN'S ACTIVITY.
 CC -|- SIMILARITY: Belongs to the serpin family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L19684; AAA59454.1; --
 CC EMBL; L28101; AAC41706.1; --
 CC EMBL; BC014992; AAH14992.1; --
 CC PIR; A49518; A49518.
 CC HSSP; POS154; 1PAI
 CC Genew; HGNC:8948; SERPINA4.
 CC MIM; 147935; --
 CC GO; GO:0005515; F:protein binding; TAS.
 CC GO; GO:0004868; F:serpin; TAS.
 CC InterPro; IPR000215; Serpin.
 CC Pfam; PF00079; serpin; 1.
 CC SMART; SM00093; SERPIN; 1.
 CC PROSITE; PS00284; SERPIN; 1.
 KW Serpin; Serine protease inhibitor; Plasma; Glycoprotein; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 427
 FT ACT SITE 388 389
 FT CARBOHYD 33 33
 FT CARBOHYD 108 108
 FT CARBOHYD 157 157
 FT CARBOHYD 238 238
 FT CONFLICT 382 382
 FT SEQUENCE 427 AA; 48556 MW; 3DBBE7AF956D4DAC CRC64;
 SQ
 Query Match 78.6%; Score 33; DB 1; Length 427;
 Best Local Similarity 66.7%; Pred. No. 7.6;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRIKASKSF 9
 DB 355 QKLEASKSF 363
 RESULT 5
 RL10 MYCPE
 ID RL10 MYCPE STANDARD; PRT; 161 AA.
 AC Q8EVJ0;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE 50S ribosomal protein L10.

```

DR PIR: S32368; S32368.
DR InterPro: IPR000744; NSF attach.
DR InterPro: IPR008941; TPR-like.
DR Pfam: PF02071; NSF; 2.
DR PRINTS: PR00448; NSFATTACHMNT.
DR Transports: Protein transport; Endoplasmic reticulum; Golgi stack.
SQ SEQUENCE 298 AA; 33555 MW; CDA8912B93A6C27E CRC64;

Query Match 73.8%; Score 31; DB 1; Length 298;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9
DB 19 KRVKASHSF 27

RESULT 8
SNAB_HUMAN STANDARD; PRT; 298 AA.
AC Q9H1I5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Beta-soluble NSF attachment protein (SNAP-beta) (N-ethylmaleimide-
DE sensitive factor attachment protein, beta).
GN NABP OR SNAPB.
OS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A.G., Cowile G.J., Deadman R., Dhani P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Levaeslao M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McLeay K., McMurray A.A.,
RA Milne S.A., Misty D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Williams S.A.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilmshurst L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.,
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
CC -!- FUNCTION: Required for vesicular transport between the endoplasmic
CC reticulum and the Golgi apparatus (By similarity).
CC -!- SUBUNIT: Interacts with PRKCAPB, and disrupts the
CC interaction between GR1A2 and PRKCAPB, leading to the
CC internalization of GR1A2 (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic peripheral membrane protein (By
CC similarity).
CC -!- SIMILARITY: Belongs to the SNAP family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL096677; CAC03439.2; ALT_SEQ.
CC Genew; HGNC:15751; NABP.
CC InterPro: IPR000744; NSF attach.
CC InterPro: IPR008941; TPR-like.
CC Pfam: PF02071; NSF; 2.
CC PRINTS: PR00448; NSFATTACHMNT.
CC Transports: Protein transport; Endoplasmic reticulum; Golgi stack.
SQ SEQUENCE 298 AA; 33557 MW; 5B7BE0F584BBD83 CRC64;

Query Match 73.8%; Score 31; DB 1; Length 298;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9
DB 19 KRVKASHSF 27

RESULT 9
TRIO_HUMAN STANDARD; PRT; 3038 AA.
AC O75962; Q13458;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Triple functional domain protein (TPPRF interacting protein).
GN TRIO.
OS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Fibroblast;
RX MEDLINE=96224308; PubMed=8643598;
RA Debant A., Serra-Pages C., Seipel K., O'Brien S., Tang M., Park S.-H.,
RA Streuli M.;
RT "The multidomain protein Trio binds the LAR transmembrane tyrosine
RT phosphatase, contains a protein kinase domain, and has separate rac-
RT specific and rho-specific guanine nucleotide exchange factor
RT domains."
RL Proc. Natl. Acad. Sci. U.S.A. 93:5466-5471(1996).
RN [2]
SEQUENCE FROM N.A. (ISOFORM 1).
RA Streuli M.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
STRUCTURE BY NMR OF 1227-1407, AND MUTAGENESIS.
RX MEDLINE=99005194; PubMed=9790533;
RA Liu X., Wang H., Eberstadt M., Schnuchel A., Olejniczak E.T.,
RA Meadows R.P., Schkeryantz J.W., Janewick D.A., Harlan J.E.,
RA Harris E.A.S., Staunton D.B., Feak S.W.;
RT "NMR structure and mutagenesis of the N-terminal Dbl homology domain
RT of the nucleotide exchange factor Trio."
RL Cell 95:269-277(1998).
RN [4]
CHARACTERIZATION.
RX MEDLINE=99276567; PubMed=10341202;
RA Seipel K., Medley Q.G., Kederbush N.L., Zhang X.A., O'Brien S.P.,
RA Serra-Pages C., Hemler M.E., Streuli M.;
RT "Trio amino-terminal guanine nucleotide exchange factor domain
RT expression promotes actin cytoskeleton reorganization, cell migration
RT and anchorage-independent cell growth."
RL J. Cell Sci. 112:1825-1834(1999).
CC -!- FUNCTION: PROMOTES THE EXCHANGE OF GDP BY GTP. TOGETHER WITH
CC LEUCOCYTE ANTIGEN-RELATED (LAR) PROTEIN, IT COULD PLAY A ROLE IN
CC COORDINATING CELL-MATRIX AND CYTOSKELETAL REARRANGEMENTS NECESSARY
CC FOR CELL MIGRATION AND CELL GROWTH.

```

GN BET1 OR SLV12 OR YIL004C OR YTA4C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN SEQUENCE FROM N.A.
RP MEDLINE=91117254; PubMed=1990290;
RX DAScher C., Ossig R., Gallwitz D., Schmitt H.D.;
RA "Identification and structure of four yeast genes (SLV) that are able
RT to suppress the functional loss of YPT1, a member of the RAS
RT superfamily.";
RL Mol. Cell. Biol. 11:872-885(1991).
RN [2].
RN SEQUENCE FROM N.A., AND FUNCTION.
RX MEDLINE=93010952; PubMed=1396561;
RA Newman A.P., Grosssch M.E., Ferro-Novick S.;
RT "Bos1p, a membrane protein required for ER to Golgi transport in
RT yeast, co-purifies with the carrier vesicles and with Bet1p and the
RT ER membrane.";
RL EMBO J. 11:3609-3617(1992).
RN [3].
RN SEQUENCE FROM N.A.
RP STRAIN=9288c;
RX MEDLINE=95282515; PubMed=7762303;
RA Voss H., Tamames J., Teodoru C., Valencia A., Sensen C., Wiemann S.,
RA Schwager C., Zimmermann J., Sander C., Ansoorge W.;
RT "Nucleotide sequence and analysis of the centromeric region of yeast
RT chromosome IX.";
RL Yeast 11:61-78(1995).
RN [4].
RN SEQUENCE FROM N.A.
RP STRAIN=9288c / AB972;
RX MEDLINE=97313266; PubMed=9169870;
RA Churcher C.M., Bowman S., Badcock K., Bankier A., Brown D.,
RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
RA Harris D.E., Hornsby T., Hunt S., Jagels K., Jones M., Lye G.,
RA Moulé S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,
RA Skelton J., Smith V., Walsh S., Whitehead S., Barrett B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX.";
RL Nature 387:84-87(1997).
CC -!- FUNCTION: Required for vesicular transport from the ER to the
CC Golgi complex. Functions as a SNARE associated with ER-derived
CC vesicles.
CC -!- SUBCELLULAR LOCATION: Type IV membrane protein (Probable).
CC -!- SIMILARITY: Belongs to the BET1 family.
CC -!- SIMILARITY: Contains 1 t-SNARE coiled-coil homology domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z38113; CA886247.1; -
DR EMBL; X54237; CA338143.1; -
DR EMBL; X79743; CA338096.1; -
DR F01; C39610; C39610.
DR GeneOnline; 139541; -
DR SGD; S0001266; BET1.
DR GO; GO:000138; C:COPII-coated vesicle; IDA.
DR GO; GO:0005789; C:endoplasmic reticulum membrane; IDA.
DR GO; GO:0005485; F:v-SNARE activity; IMP.
DR GO; GO:0005888; P:ER to Golgi transport; IMP.
DR GO; GO:0006890; P:retrograde (Golgi to ER) transport; IDA.
DR InterPro; IPR000277; t-SNARE.
DR Pfam; PF05739; SNARE; 1.
DR SMART; SM00397; t-SNARE; 1.
DR PROSITE; PS00192; t-SNARE; 1.
DR Transport; Protein transport; Endoplasmic reticulum; Golgi stack;
KW Transmembrane; Coiled coil.

FT DOMAIN 1 117 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 118 141 ANCHOR FOR TYPE IV MEMBRANE PROTEIN
FT DOMAIN 142 142 (POTENTIAL).
FT DOMAIN 52 114 VESICULAR (POTENTIAL).
FT SEQUENCE 142 AA; 15720 MW; DC1BD4FFE67E2217 CRC64;
[1]
Query Match 71.4%; Score 30; DB 1; Length 142;
Best Local Similarity 87.5%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 QRIKASKS 8
| | | | |
DB 67 QRIKALKS 74

RESULT 11
YB3C SCHPO
ID YB3C SCHPO STANDARD; PRT; 517 AA.
AC O14342;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C2F12.12c in chromosome II.
GN SPBC2F12.12C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1].
RN SEQUENCE FROM N.A.
RP STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Soutos J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moulé S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Ruster S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voiclaert G., Aert R., Robben J., Grymprez B.,
RA Welter J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Mostl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z97211; CAB10159.1; -
DR F01; T40129; T40129.
DR GeneDB Spombe; SPBC2F12.12C; -
KW Hypothetical protein.

```

DR EMBL; L04938; AAA25204.1; -.
DR EMBL; U09553; AAB00538.1; -.
DR PIR; A47098; A47098.
DR MEROPS; M13.004; -.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000718; Peptidase_M13.
DR InterPro; IPR008753; Peptidase_M13_N.
DR PIR; PF01431; Peptidase_M13; 1.
DR PIR; PF05649; Peptidase_M13_N; 1.
DR PRINTS; PR00786; NEPRILYSIN.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Zinc.
FT INIT MET 0
FT METAL 474 474 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT SITE 475 475 ZINC (CATALYTIC).
FT METAL 478 478 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 534 534 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT SITE 538 538 ZINC (CATALYTIC) (BY SIMILARITY).
SQ SEQUENCE 626 AA; 71396 MW; DDFED4FE47FEEOC CRC64;

Query Match 71.4%; Score 30; DB 1; Length 626;
Best Local Similarity 75.0%; Pred. No. 55;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RIKASKSF 9
DB 569 RYKASKSF 576

RESULT 14
PKP1_BOVIN
ID PKP1_BOVIN STANDARD; PRT; 727 AA.
AC Q28161;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Plakophilin 1 (Band-6-protein) (B6P).
GN PKP1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95196971; PubMed=7890138;
RA Heid H.W., Schmidt A., Zimbelmann R., Schaefer S.,
RA Winter-Sinanowski S., Stump S., Keith M., Fige U., Schnolzer M.,
RA Franke W.W.;
RT "Cell type-specific desmosomal plaque proteins of the plakoglobin
RT family: plakophilin 1 (band 6 protein).";
RL Differentiation 58:113-131(1994).
CC -!- FUNCTION: SEEMS TO PLAY A ROLE IN JUNCTIONAL PLAQUES (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Nuclear and associated with desmosomes (By
CC similarity).
CC -!- SIMILARITY: Belongs to the beta-catenin family.
CC -!- SIMILARITY: Contains 8 ARM repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z37975; CAA86029.1; -.
CC DR GO; GO:0030057; C:desmosome; ISS.
CC DR GO; GO:0005634; C:nucleus; ISS.
CC DR GO; GO:0019215; F:intermediate filament binding; ISS.
CC DR GO; GO:0030280; F:structural constituent of epidermis; ISS.
CC DR GO; GO:0007155; P:cell adhesion; ISS.

QY 2 RIKASKSF 9
DB 569 RYKASKSF 576

RESULT 15
PKP1_MOUSE
ID PKP1_MOUSE STANDARD; PRT; 728 AA.
AC P97350;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Plakophilin 1.
GN PKP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6; TISSUE=Skin; Franke W.W.;
RA Nimmich V.; Hunziker A.H.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: SEEMS TO PLAY A ROLE IN JUNCTIONAL PLAQUES (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Nuclear and associated with desmosomes (By
CC similarity).
CC -!- SIMILARITY: Belongs to the beta-catenin family.
CC -!- SIMILARITY: Contains 7 ARM repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y07941; CAA69240.1; -.
CC DR MGD; MGI:1328359; Pkpl.
CC DR GO; GO:0030057; C:desmosome; ISS.
CC DR GO; GO:0005634; C:nucleus; ISS.
CC DR GO; GO:0019215; F:intermediate filament binding; ISS.
CC DR GO; GO:0030280; F:structural constituent of epidermis; ISS.
CC DR GO; GO:0007155; P:cell adhesion; ISS.
CC DR GO; GO:0007165; P:signal transduction; ISS.
CC DR InterPro; IPR008938; ARM.
CC DR InterPro; IPR000225; Armadillo.
CC DR Pfam; PF00514; Armadillo_seg_4.
CC DR SMART; SM00185; ARM; 6.
CC DR PROSITE; PS50176; ARM_REPEAT; 3.

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 17:20:41 ; Search time 30.3333 Seconds
(without alignments)
93.615 Million cell updates/sec

Title: US-09-905-083-99

Perfect score: 42

Sequence: 1 QRIKASKSF 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25:**

1: sp_archaea:**

2: sp_bacteria:**

3: sp_fungi:**

4: sp_human:**

5: sp_invertebrate:**

6: sp_mammal:**

7: sp_mbc:**

8: sp_organelle:**

9: sp_phage:**

10: sp_plant:**

11: sp_rodent:**

12: sp_virus:**

13: sp_vertebrate:**

14: sp_unclassified:**

15: sp_virus:**

16: sp_bacteriap:**

17: sp_archaea:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	100.0	181	Q8NFV7	Q8NFV7 homo sapien
2	42	100.0	253	Q8NSN9	Q8NSN9 homo sapien
3	36	85.7	234	Q9R048	Q9R048 mus musculu
4	36	85.7	249	Q91VE3	Q91VE3 mus musculu
5	34	81.0	294	Q88Q37	Q88Q37 pseudomonas
6	34	81.0	683	Q82XB0	Q82XB0 nitrosomona
7	33	78.6	445	Q86TR9	Q86TR9 homo sapien
8	33	78.6	671	Q8EUP9	Q8EUP9 mycoplasma
9	32	76.2	274	Q8U023	Q8U023 leishmania
10	32	76.2	467	Q8IBN4	Q8IBN4 plasmodium
11	32	76.2	474	Q7VRX5	Q7VRX5 candidatus
12	32	76.2	490	Q8SAQ6	Q8SAQ6 chlamydomon
13	32	76.2	684	Q816K2	Q816K2 holotrichia
14	32	76.2	895	Q9YAU1	Q9YAU1 aeropyrum p
15	31	73.8	177	Q8TB05	Q8TB05 homo sapien
16	31	73.8	248	Q937X0	Q937X0 sphingomona

17	31	73.8	254	17	Q975X7	Q975X7 sulfobolus
18	31	73.8	276	4	Q8N8N1	Q8N8N1 homo sapien
19	31	73.8	302	3	Q9UTE6	Q9UTE6 schizosacch
20	31	73.8	362	16	Q9PHQ2	Q9PHQ2 campylobact
21	31	73.8	443	5	Q8TA05	Q8TA05 acrasis ros
22	31	73.8	895	16	Q8PPV7	Q8PPV7 xanthomonas
23	31	73.8	895	16	Q8P4T1	Q8P4T1 xanthomonas
24	31	73.8	1627	4	Q81WKS	Q81WKS homo sapien
25	31	73.8	2503	5	Q72017	Q72017 penaeus van
26	30	71.4	80	16	Q8XVX1	Q8XVX1 raietonia s
27	30	71.4	103	17	Q8U1I5	Q8U1I5 pyrococcus
28	30	71.4	211	16	Q88RJ0	Q88RJ0 pseudomonas
29	30	71.4	321	16	Q84847	Q84847 chlamydia t
30	30	71.4	328	16	Q81RD6	Q81RD6 bacillus an
31	30	71.4	328	16	Q81E81	Q81E81 bacillus ce
32	30	71.4	331	16	Q97JF2	Q97JF2 clostridium
33	30	71.4	427	10	Q9M035	Q9M035 arabidopsis
34	30	71.4	428	16	Q823E3	Q823E3 chlamydophi
35	30	71.4	475	5	Q81EQ6	Q81EQ6 plasmodium
36	30	71.4	478	16	Q926K2	Q926K2 listeria in
37	30	71.4	527	5	Q86K80	Q86K80 dictyosteli
38	30	71.4	582	16	Q8XP65	Q8XP65 clostridium
39	30	71.4	583	10	Q9FM26	Q9FM26 arabidopsis
40	30	71.4	583	10	Q93ZB5	Q93ZB5 arabidopsis
41	30	71.4	588	11	Q8CDE2	Q8CDE2 mus musculu
42	30	71.4	826	5	Q968Z5	Q968Z5 caenorhabdi
43	30	71.4	844	5	Q18372	Q18372 caenorhabdi
44	30	71.4	896	16	Q87BF1	Q87BF1 xylella fas
45	30	71.4	917	5	Q86L00	Q86L00 dictyosteli

ALIGNMENTS

RESULT 1

Q8NFV7	Q8NFV7	PRELIMINARY;	PRT;	181 AA.
ID	Q8NFV7	PRELIMINARY;	PRT;	181 AA.
AC	Q8NFV7;			
DT	01-OCT-2002 (TREMBLrel. 22, Created)			
DT	01-OCT-2002 (TREMBLrel. 22, Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DE	Kallikrein 7 short variant protein.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Ovarian carcinoma;			
RA	Dong Y., Kaushal A., Clements J.A.;			
RT	"Human Kallikrein 7 (KLK7) short variant mRNA from ovarian carcinoma.";			
RL	Submitted (AUG-2001) to the EMBL/GenBank/DBSJ databases.			
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			
DR	EMBL; AF411215; AAN03663.1; -			
DR	GO; GO:0004263; F:chymotrypsin activity; IEA.			
DR	GO; GO:0008233; F:peptidase activity; IEA.			
DR	GO; GO:0004295; F:trypsin activity; IEA.			
DR	GO; GO:0008508; P:proteolysis and peptidolysis; IEA.			
DR	InterPro; IPR009003; Cys_Ser_trypsin.			
DR	InterPro; IPR001254; Peptidase_S1.			
DR	Pfam; PF00089; trypsin; 1.			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	SMART; SM00020; Tryp_SPC; 1.			
DR	PROSITE; PS02440; TRYPSIN_DOM; 1.			
DR	PROSITE; PS00135; TRYPSIN_SER; 1.			
KW	Hydrolase; Protease; Serine protease.			
SQ	SEQUENCE 181 AA; 19887 MW; 8628A03B80C2D78 CRC64;			

Query Match 100.0%; Score 42; DB 4; Length 181;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

DR EMBL; BC027823; AAH27823.1; -
DR EMBL; AK029477; BAC26467.1; -
DR EMBL; AK077406; BAC36787.1; -
DR HSSP; P00761; IAN1.
DR MGD; MGI:1346336; Kik7.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PROSITE; PS0240; TRYPsin DOM; 1.
DR PROSITE; PS00134; TRYPsin HIS; 1.
DR PROSITE; PS00135; TRYPsin SR; 1.
DR Hydrolase; Process; Serine protease.
KW SEQUENCE 249 AA; 27257 MW; 0D4E380F12D14F87 CRC64;

Query Match 85.7%; Score 36; DB 11; Length 249;
Best Local Similarity 77.8%; Pred. NO. 8.1;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9
Db 87 QRIKATKSF 95

RESULT 5
Q88Q37 PRELIMINARY; PRT; 294 AA.
AC Q88Q37 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Transcriptional regulator AmpR, putative.
GN PP0661.
OS Pseudomonas putida (strain KT2440)
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Beanan M., DeBoy R.F., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzes A.,
RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
RA Lauber J., Stjepandic D., Hobeisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808 (2002).
DR EMBL; AE016776; AAN66286.1; -
DR TIGR; TP0661; -
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000847; HTH_LysR.
DR InterPro; IPR005119; LysR_subst.
DR Pfam; PF00126; HTH 1; 1.
DR Pfam; PF03466; LysR_substrate; 1.
DR PRINTS; PRO0039; HTHLYSR.
DR PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
KW Complete proteome.
SQ SEQUENCE 294 AA; 32995 MW; 33695F24D26560EF CRC64;

Query Match 81.0%; Score 34; DB 16; Length 294;
Best Local Similarity 77.8%; Pred. NO. 27;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 QRIKASKSF 9
Db 259 QRIEASESF 267

RESULT 6
Q82XB0 PRELIMINARY; PRT; 683 AA.
AC Q82XB0 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sensory transduction histidine kinases.
GN NE0377.
OS Nitrosomonas europaea.
OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC Nitrosomonadaceae; Nitrosomonas.
OX NCBI_TaxID=915;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 19718 / IFO 14298;
RX MEDLINE=22586410; PubMed=12700255;
RA Chain P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.,
RA Hauser L., Hooper A.B., Klotz M.G., Norton J., Sayavedra-Soto L.A.,
RA Arciero D.M., Hommes N.G., Whittaker M.M., Arp D.J.;
RT "Complete genome sequence of the ammonia-oxidizing bacterium and
RT obligate chemolithoautotroph Nitrosomonas europaea.";
RL J. Bacteriol. 185:2759-2773 (2003).
DR EMBL; BX321857; CAD84288.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR003594; ATPbind ATPase.
DR InterPro; IPR004358; Bact_sens_dr_C.
DR InterPro; IPR005467; His_Kinase.
DR InterPro; IPR003661; His_kinA_N.
DR Pfam; PF02518; HATFase_c; 1.
DR Pfam; PF00512; Hiska; 1.
DR PRINTS; PRO0344; BCTRISensor.
DR SMART; SM00387; HATFase_c; 1.
DR PROSITE; PS01019; HIS_KIN; 1.
DR Kinase; Complete proteome.
SQ SEQUENCE 683 AA; 77013 MW; 6989B95EB543ED4 CRC64;

Query Match 81.0%; Score 34; DB 16; Length 683;
Best Local Similarity 88.9%; Pred. NO. 62;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QRIKASKSF 9
Db 548 QRIIASKSF 556

RESULT 7
Q86TR9 PRELIMINARY; PRT; 445 AA.
AC Q86TR9 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Human full-length cDNA 5-PRIME end of clone CS0DM009YC13 of fetal
DE liver of Homo sapiens (Human) (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tissue=Fetal liver;
RC Genoscope;

```

```
RESULT 14
Q9YAU1 ID Q9YAU1 PRELIMINARY; PRT; 895 AA.
AC Q9YAU1
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 895AA long hypothetical DNA-directed RNA polymerase subunit A'.
GN APE1853.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1.
RX MEDLINE=99310339; PubMed=10382966;
RA Kawaiabayashi Y., Hino Y., Horikawa H., Yamazaki S., Hakawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Arkai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya A., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1."
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000062; BAA80857.1; -
DR PIR; D72571; D72571.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003899; F:DNA-directed RNA polymerase activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR006592; RNA_polA_N.
DR InterPro; IPR007022; RNA_pol_A.
DR InterPro; IPR007080; RNA_pol_Rpb1_1.
DR InterPro; IPR007066; RNA_pol_Rpb1_3.
DR InterPro; IPR007083; RNA_pol_Rpb1_4.
DR InterPro; IPR007081; RNA_pol_Rpb1_5.
DR Pfam; PF04997; RNA_pol_Rpb1_1; 1.
DR Pfam; PF06623; RNA_pol_Rpb1_2; 1.
DR Pfam; PF04983; RNA_pol_Rpb1_3; 1.
DR Pfam; PF05000; RNA_pol_Rpb1_4; 1.
DR Pfam; PF04998; RNA_pol_Rpb1_5; 1.
DR SMART; SM00663; RPOA_N; 1.
KW DNA-directed RNA polymerase; Complete proteome.
SQ SEQUENCE 895 AA; 101740 MW; 1E01EC06B4BD0AAE CRC64;

Query Match 76.2%; Score 32; DB 17; Length 895;
Best Local Similarity 75.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRIKASKS 8
DB 885 QRVKASKA 892

RESULT 15
Q8TB05 ID Q8TB05 PRELIMINARY; PRT; 177 AA.
AC Q8TB05
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to LOC146557.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; BC025327; AAH25327.1; -
SQ SEQUENCE 177 AA; 18954 MW; 7CCC710752C42712 CRC64;

Query Match 73.8%; Score 31; DB 4; Length 177;
Best Local Similarity 75.0%; Pred. No. 77;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 RIKASKSF 9
DB 84 RLKASESF 91

Search completed: March 1, 2004, 17:34:56
Job time : 32.3333 secs
```